

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/82, 15/54	A2	(11) International Publication Number: WO 98/46776 (43) International Publication Date: 22 October 1998 (22.10.98)
(21) International Application Number: PCT/US98/07114 (22) International Filing Date: 9 April 1998 (09.04.98) (30) Priority Data: 60/041,815 11 April 1997 (11.04.97) US (71) Applicant (for all designated States except US): CALGENE LLC [US/US]; 1920 Fifth Street, Davis, CA 95616 (US). (72) Inventor; and (75) Inventor/Applicant (for US only): DEHESH, Katayoon [US/US]; 521 Crownpointe Circle, Vacaville, CA 95687 (US). (74) Agent: SCHWEDLER, Carl, J.; Calgene LLC, 1920 Fifth Street, Davis, CA 95616 (US).		(81) Designated States: AU, BR, CA, JP, KR, MX, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: PLANT FATTY ACID SYNTHASES AND USE IN IMPROVED METHODS FOR PRODUCTION OF MEDIUM-CHAIN FATTY ACIDS (57) Abstract <p>By this invention, compositions and methods of use related to β-ketoacyl-ACP synthase of special interest are synthases obtainable from <i>Cuphea</i> species. Amino acid and nucleic acid for synthase protein factors are provided, as well as methods to utilize such sequences in constructs for production of genetically engineered plants having altered fatty acid compositions. Of particular interest is the expression of synthase protein factors in conjunction with expression of plant medium-chain acyl-ACP thioesterases for production of increased levels and/or modified ratios of medium-chain fatty acids in oils of transgenic plant seeds.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

PLANT FATTY ACID SYNTHASES AND USE IN IMPROVED METHODS FOR
PRODUCTION OF MEDIUM-CHAIN FATTY ACIDS

5

INTRODUCTION

Field of Invention

The present invention is directed to genes encoding
10 plant fatty acid synthase enzymes relevant to fatty acid
synthesis in plants, and to methods of using such genes in
combination with genes encoding plant medium-chain
preferring thioesterase proteins. Such uses provide a
method to increase the levels of medium-chain fatty acids
15 that may be produced in seed oils of transgenic plants.

Background

Higher plants synthesize fatty acids via a common
metabolic pathway. In developing seeds, where fatty acids
20 attached to triglycerides are stored as a source of energy
for further germination, the fatty acid synthesis pathway is
located in the plastids. The first step is the formation of
acetyl-ACP (acyl carrier protein) from acetyl-CoA and ACP
catalyzed by a short chain preferring condensing enzyme, β -
25 ketoacyl-ACP synthase (KAS) III. Elongation of acetyl-ACP
to 16- and 18- carbon fatty acids involves the cyclical
action of the following sequence of reactions: condensation
with a two-carbon unit from malonyl-ACP to form a longer β -
ketoacyl-ACP (β -ketoacyl-ACP synthase), reduction of the

keto-function to an alcohol (β -ketoacyl-ACP reductase), dehydration to form an enoyl-ACP (β -hydroxyacyl-ACP dehydrase), and finally reduction of the enoyl-ACP to form the elongated saturated acyl-ACP (enoyl-ACP reductase). β -ketoacyl-ACP synthase I (KAS I), is primarily responsible for elongation up to palmitoyl-ACP (C16:0), whereas β -ketoacyl-ACP synthase II (KAS II) is predominantly responsible for the final elongation to stearoyl-ACP (C18:0).

Genes encoding peptide components of β -ketoacyl-ACP synthases I and II have been cloned from a number of higher plant species, including castor (*Ricinus communis*) and *Brassica* species (USPN 5,510,255). KAS I activity was associated with a single synthase protein factor having an approximate molecular weight of 50 kD (synthase factor B) and KAS II activity was associated with a combination of two synthase protein factors, the 50 kD synthase factor B and a 46 kD protein designated synthase factor A. Cloning and sequence of a plant gene encoding a KAS III protein has been reported by Tai and Jaworski (*Plant Physiol.* (1993) 103:1361-1367).

The end products of plant fatty acid synthetase activities are usually 16- and 18-carbon fatty acids. There are, however, several plant families that store large amounts of 8- to 14-carbon (medium-chain) fatty acids in their oilseeds. Recent studies with *Umbellularia californica* (California bay), a plant that produces seed oil rich in lauric acid (12:0), have demonstrated the existence of a medium-chain-specific isozyme of acyl-ACP thioesterase

in the seed plastids. Subsequent purification of the 12:0-ACP thioesterase from *Umbellularia californica* led to the cloning of a thioesterase cDNA which was expressed in seeds of *Arabidopsis* and *Brassica* resulting in a substantial accumulation of lauric acid in the triglyceride pools of these transgenic seeds (USPN 5,512,482). These results and subsequent studies with medium-chain thioesterases from other plant species have confirmed the chain-length-determining role of acyl-ACP thioesterases during de novo fatty acid biosynthesis (T. Voelker (1996) *Genetic Engineering*, Ed. J. K. Setlow, Vol. 18, pgs. 111-133).

DESCRIPTION OF THE FIGURES

- Figure 1. DNA and translated amino acid sequence of *Cuphea hookeriana* KAS factor B clone chKAS B-2 are provided.
- Figure 2. DNA and translated amino acid sequence of *Cuphea hookeriana* KAS factor B clone chKAS B-31-7 are provided.
- Figure 3. DNA and translated amino acid sequence of *Cuphea hookeriana* KAS factor A clone chKAS A-2-7 are provided.
- Figure 4. DNA and translated amino acid sequence of *Cuphea hookeriana* KAS factor A clone chKAS A-1-6 are provided.
- Figure 5. DNA and translated amino acid sequence of *Cuphea pullcherrima* KAS factor B clone cpuKAS B/7-8 are provided.
- Figure 6. DNA and translated amino acid sequence of *Cuphea pullcherrima* KAS factor B clone cpuKAS B/8-7A are provided.
- Figure 7. DNA and translated amino acid sequence of *Cuphea pullcherrima* KAS factor A clone cpuKAS A/p7-6A are provided.
- Figure 8. Preliminary DNA sequence of *Cuphea pullcherrima* KAS factor A clone cpuKAS A/p8-9A is provided.

Figure 9. DNA and translated amino acid sequence of *Cuphea hookeriana* KASIII clone chKASIII-27 are provided.

Figure 10. The activity profile for purified cpuKAS B/8-7A using various acyl-ACP substrates is provided.

5 Figure 11. The activity profile for purified chKAS A-2-7 and chKAS A-1-6 using various acyl-ACP substrates is provided.

Figure 12. The activity profile for purified castor KAS factor A using various acyl-ACP substrates is provided.

10 Figure 13. The activity profile for purified castor KAS factor B using various acyl-ACP substrates is provided.

Figure 14. A graph showing the number of plants arranged according to C8:0 content for transgenic plants containing CpFatB1 versus transgenic plants containing CpFatB1 + chKAS
15 A-2-7 is provided.

Figure 15. Graphs showing the %C10/%C8 ratios in transgenic plants containing ChFatB2 (4804-22-357) and in plants resulting from crosses between 4804-22-357 and 5401-9 (chKAS A-2-7 plants) are provided.

20 Figure 16. Graphs showing the %C10 + %C8 contents in transgenic plants containing ChFatB2 (4804-22-357) and in plants resulting from crosses between 4804-22-357 and 5401-9 (chKAS A-2-7 plants) are provided.

Figure 17. Graphs showing the %C10/%C8 ratios in transgenic
25 plants containing ChFatB2 (4804-22-357) and in plants resulting from crosses between 4804-22-357 and 5413-17 (chKAS A-2-7 + CpFatB1 plants) are provided.

Figure 18. Graphs showing the %C10 + %C8 contents in transgenic plants containing ChFatB2 (4804-22-357) and in

plants resulting from crosses between 4804-22-357 and 5413-17 (chKAS A-2-7 + CpFatB1 plants) are provided.

Figure 19. Graphs showing the %C12:0 in transgenic plants containing Uc FatB1 (LA86DH186) and in plants resulting from crosses with wild type (X WT) and with lines expressing Ch KAS A-2-7.

Figure 20. Graph showing the relative proportions of C12:0 and C14:0 fatty acids in the seeds of transgenic plants containing Uc FatB1 (LA86DH186) and in plants resulting from crosses with wild type (X WT) and with lines expressing Ch KAS A-2-7.

Figure 21. Graphs showing the %C18:0 in transgenic plants containing Garm FatB1 (5266) and in seeds of plants resulting from crosses with wild type (X WT) and with lines expressing Ch KAS A-2-7.

Figure 22. The activity profile of Ch KAS A in protein extracts from transgenic plants containing Ch KAS A-2-7. Extracts were pretreated with the indicated concentrations of cerulenin.

SUMMARY OF THE INVENTION

By this invention, compositions and methods of use related to β -ketoacyl-ACP synthase (KAS) are provided. Also of interest are methods and compositions of amino acid and nucleic acid sequences related to biologically active plant synthase(s).

In particular, genes encoding KAS protein factors A and B from *Cuphea* species are provided. The KAS genes are of interest for use in a variety of applications, and may be

used to provide synthase I and/or synthase II activities in transformed host cells, including bacterial cells, such as *E. coli*, and plant cells. Synthase activities are distinguished by the preferential activity towards longer and shorter acyl-ACPs as well as by the sensitivity towards the KAS specific inhibitor, cerulenin. Synthase protein preparations having preferential activity towards medium chain length acyl-ACPs are synthase I-type or KAS I. The KAS I class is sensitive to inhibition by cerulenin at concentrations as low as 1 μ M. Synthases having preferential activity towards longer chain length acyl-ACPs are synthase II-type or KAS II. The KAS enzymes of the II-type are also sensitive to cerulenin, but at higher concentrations (50 μ M). Synthase III-type enzymes have preferential activity towards short chain length acyl-ACPs and are insensitive to cerulenin inhibition.

Nucleic acid sequences encoding a synthase protein may be employed in nucleic acid constructs to modulate the amount of synthase activity present in the host cell, especially the relative amounts of synthase I-type, synthase II-type and synthase III-type activity when the host cell is a plant host cell. In addition, nucleic acid constructs may be designed to decrease expression of endogenous synthase in a plant cell as well. One example is the use of an anti-sense synthase sequence under the control of a promoter capable of expression in at least those plant cells which normally produce the enzyme.

Of particular interest in the present invention is the coordinate expression of a synthase protein with the

expression of thioesterase proteins. For example, coordinated expression of synthase factor A and a medium-chain thioesterase provides a method for increasing the level of medium-chain fatty acids which may be harvested from transgenic plant seeds. Furthermore, coordinated expression of a synthase factor A gene with plant medium-chain thioesterase proteins also provides a method by which the ratios of various medium-chain fatty acids produced in a transgenic plant may be modified. For example, by expression of a synthase factor A, it is possible to increase the ratio of C10/C8 fatty acids which are produced in plant seed oils as the result of expression of a thioesterase having activity on C8 and C10 fatty acids.

DETAILED DESCRIPTION OF THE INVENTION

A plant synthase factor protein of this invention includes a sequence of amino acids or polypeptide which is required for catalyzation of a condensation reaction between an acyl-ACP having a chain length of C₂ to C₁₆ and malonyl-ACP in a plant host cell. A particular plant synthase factor protein may be capable of catalyzing a synthase reaction in a plant host cell (for example as a monomer or homodimer) or may be one component of a multiple peptide enzyme which is capable of catalyzing a synthase reaction in a plant host cell, i.e. one peptide of a heterodimer.

Synthase I (KAS I) demonstrates preferential activity towards acyl-ACPs having shorter carbon chains, C₂-C₁₄ and is sensitive to inhibition by cerulenin at concentrations of 1 μ M. Synthase II (KAS II) demonstrates preferential

activity towards acyl-ACPs having longer carbon chains, C₁₄-C₁₆, and is inhibited by concentrations of cerulenin (50μM). Synthase III demonstrates preferential activity towards acyl-CoAs having very short carbon chains, C₂ to C₆, and is
5 insensitive to inhibition by cerulenin.

Synthase factors A and B, and synthase III proteins obtained from medium-chain fatty acid producing plant species of the genus *Cuphea* are described herein. As described in the following Examples, synthase A from *C. hookeriana* is naturally expressed at a high level and only
10 in the seeds. *C. hookeriana* synthase B is expressed at low levels in all tissues examined. Expression of synthase A and synthase B factors in *E. coli* and purification of the resulting proteins is employed to determine activity of the
15 various synthase factors. Results of these analyses indicate that synthase factor A from *Cuphea hookeriana* has the greatest activity on 6:0-ACP substrates, whereas synthase factor B from *Cuphea pullcherrima* has greatest activity on 14:0-ACP. Similar studies with synthase factors
20 A and B from castor demonstrate similar activity profiles between the factor B synthase proteins from *Cuphea* and castor. The synthase A clone from castor, however, demonstrates a preference for 14:0-ACP substrate.

Expression of a *Cuphea hookeriana* KAS A protein in
25 transgenic plant seeds which normally do not produce medium-chain fatty acids does not result in any detectable modification of the fatty acid types and contents produced in such seeds. However, when *Cuphea hookeriana* KAS A protein is expressed in conjunction with expression of a

medium-chain acyl-ACP thioesterase capable of providing for production of C8 and C10 fatty acids in plant seed oils, increases in the levels of medium-chain fatty acids over the levels obtainable by expression of the medium-chain

5 thioesterase alone are observed. In addition, where significant amounts of C8 and C10 fatty acids are produced as the result of medium-chain thioesterase expression, co-expression of a *Cuphea* KAS A protein also results in an alteration of the proportion of the C8 and C10 fatty acids

10 that are obtained. For example, an increased proportion of C10 fatty acids may be obtained by co-expression of *Cuphea hookeriana* ChFatB2 thioesterase and a chKAS A synthase factor proteins.

Furthermore, when *Cuphea hookeriana* KAS A protein is

15 expressed in conjunction with expression of a medium-chain acyl-ACP thioesterase capable of providing for production of C12 fatty acids in plant seed oils, increases in the levels of medium-chain fatty acids over the levels obtainable by expression of the medium-chain thioesterase alone are also

20 observed. In addition, where significant amounts of C12 and C14 fatty acids are produced as the result of medium-chain thioesterase expression, co-expression of a *Cuphea* KAS A protein also results in an alteration of the proportion of the C12 and C14 fatty acids that are obtained. For example,

25 an increased proportion of C12 fatty acids may be obtained by co-expression of *Uc* FatB1 thioesterase and a chKAS A synthase factor proteins.

However, when *Cuphea hookeriana* KAS A protein is expressed in conjunction with the expression of a long-chain

acyl-ACP thioesterase capable of providing for production of C18 and C18:1 fatty acids in plant seed oils, no effect on the production of long chain fatty acids was observed.

Furthermore, when plants transformed to express a long chain
5 acyl-ACP thioesterase from mangosteen (*GarmFatA1*, U.S.

Patent Application No. 08/440,845), which preferentially hydrolyzes C18:0 and C18:1 fatty acyl-ACPs, are crossed with nontransformed control plants, a significant reduction in the levels of C18:0 is obtained. Similar reductions are also
10 observed in the levels of C18:0 in the seeds of plants resulting from crosses between plants transformed to express the *GarmFatA1* and plants expressing the *Cuphea hookeriana* KAS A protein.

Thus, the instant invention provides methods of
15 increasing and/or altering the medium-chain fatty acid compositions in transgenic plant seed oils by co-expression of medium-chain acyl-ACP thioesterases with synthase factor proteins. Furthermore, various combinations of synthase factors and medium-chain thioesterases may be achieved
20 depending upon the particular fatty acids desired. For example, for increased production of C14 fatty acids, synthase protein factors may be expressed in combination with a C14 thioesterase, for example from *Cuphea palustris* or nutmeg may be employed (WO 96/23892). In addition,
25 thioesterase expression may be combined with a number of different synthase factor proteins for additional effects on medium-chain fatty acid composition.

Synthases of use in the present invention include modified amino acid sequences, such as sequences which have

been mutated, truncated, increased and the like, as well as such sequences which are partially or wholly artificially synthesized. The synthase protein encoding sequences provided herein may be employed in probes for further screening or used in genetic engineering constructs for transcription or transcription and translation in host cells, especially plant host cells. One skilled in the art will readily recognize that antibody preparations, nucleic acid probes (DNA and RNA) and the like may be prepared and used to screen and recover synthases and/or synthase nucleic acid sequences from other sources. Typically, a homologously related nucleic acid sequence will show at least about 60% homology, and more preferably at least about 70% homology, between the *R. communis* synthase and the given plant synthase of interest, excluding any deletions which may be present. Homology is determined upon comparison of sequence information, nucleic acid or amino acid, or through hybridization reactions.

Recombinant constructs containing a nucleic acid sequence encoding a synthase protein factor or nucleic acid sequences encoding a synthase protein factor and a medium-chain acyl-ACP thioesterase may be prepared by methods well known in the art. Constructs may be designed to produce synthase in either prokaryotic or eukaryotic cells. The increased expression of a synthase in a plant cell, particularly in conjunction with expression of medium-chain thioesterases, or decreasing the amount of endogenous synthase observed in plant cells are of special interest.

Synthase protein factors may be used, alone or in combination, to catalyze the elongating condensation reactions of fatty acid synthesis depending upon the desired result. For example, rate influencing synthase activity may
5 reside in synthase I-type, synthase II-type, synthase III-type or in a combination of these enzymes. Furthermore, synthase activities may rely on a combination of the various synthase factors described herein.

Constructs which contain elements to provide the
10 transcription and translation of a nucleic acid sequence of interest in a host cell are "expression cassettes". Depending upon the host, the regulatory regions will vary, including regions from structural genes from viruses, plasmid or chromosomal genes, or the like. For expression
15 in prokaryotic or eukaryotic microorganisms, particularly unicellular hosts, a wide variety of constitutive or regulatable promoters may be employed. Among transcriptional initiation regions which have been described are regions from bacterial and yeast hosts, such as *E. coli*,
20 *B. subtilis*, *Saccharomyces cerevisiae*, including genes such as β -galactosidase, T7 polymerase, trp-lac (tac), trp E and the like.

An expression cassette for expression of synthase in a plant cell will include, in the 5' to 3' direction of
25 transcription, a transcription and translation initiation control regulatory region (also known as a "promoter") functional in a plant cell, a nucleic acid sequence encoding a synthase, and a transcription termination region. Numerous transcription initiation regions are available

which provide for a wide variety of constitutive or regulatable, e.g., inducible, transcription of the desaturase structural gene. Among transcriptional initiation regions used for plants are such regions associated with cauliflower mosaic viruses (35S, 19S), and structural genes such as for nopaline synthase or mannopine synthase or napin and ACP promoters, etc. The transcription/ translation initiation regions corresponding to such structural genes are found immediately 5' upstream to the respective start codons. Thus, depending upon the intended use, different promoters may be desired.

Of special interest in this invention are the use of promoters which are capable of preferentially expressing the synthase in seed tissue, in particular, at early stages of seed oil formation. Examples of such seed-specific promoters include the region immediately 5' upstream of a napin or seed ACP genes such as described in USPN 5,420,034, desaturase genes such as described in Thompson et al (*Proc. Nat. Acad. Sci.* (1991) 88:2578-2582), or a Bce-4 gene such as described in USPN 5,530,194. Alternatively, the use of the 5' regulatory region associated with the plant synthase structural gene, i.e., the region immediately 5' upstream to a plant synthase structural gene and/or the transcription termination regions found immediately 3' downstream to the plant synthase structural gene, may often be desired. In general, promoters will be selected based upon their expression profile which may change given the particular application.

In addition, one may choose to provide for the transcription or transcription and translation of one or more other sequences of interest in concert with the expression or anti-sense of the synthase sequence, particularly medium-chain plant thioesterases such as described in USPN 5,512,482, to affect alterations in the amounts and/or composition of plant oils.

When one wishes to provide a plant transformed for the combined effect of more than one nucleic acid sequence of interest, a separate nucleic acid construct may be provided for each or the constructs may both be present on the same plant transformation construct. The constructs may be introduced into the host cells by the same or different methods, including the introduction of such a trait by crossing transgenic plants via traditional plant breeding methods, so long as the resulting product is a plant having both characteristics integrated into its genome.

Normally, included with the DNA construct will be a structural gene having the necessary regulatory regions for expression in a host and providing for selection of transformed cells. The gene may provide for resistance to a cytotoxic agent, e.g. antibiotic, heavy metal, toxin, etc., complementation providing prototrophy to an auxotrophic host, viral immunity or the like. Depending upon the number of different host species into which the expression construct or components thereof are introduced, one or more markers may be employed, where different conditions for selection are used for the different hosts.

The manner in which the DNA construct is introduced into the plant host is not critical to this invention. Any method which provides for efficient transformation may be employed. Various methods for plant cell transformation include the use of Ti- or Ri-plasmids, microinjection, electroporation, liposome fusion, DNA bombardment or the like. In many instances, it will be desirable to have the construct bordered on one or both sides by T-DNA, particularly having the left and right borders, more particularly the right border. This is particularly useful when the construct uses *A. tumefaciens* or *A. rhizogenes* as a mode for transformation, although the T-DNA borders may find use with other modes of transformation.

The expression constructs may be employed with a wide variety of plant life, particularly plant life involved in the production of vegetable oils. These plants include, but are not limited to rapeseed, peanut, sunflower, safflower, cotton, soybean, corn and oilseed palm.

For transformation of plant cells using *Agrobacterium*, explants may be combined and incubated with the transformed *Agrobacterium* for sufficient time for transformation, the bacteria killed, and the plant cells cultured in an appropriate selective medium. Once callus forms, shoot formation can be encouraged by employing the appropriate plant hormones in accordance with known methods and the shoots transferred to rooting medium for regeneration of plants. The plants may then be grown to seed and the seed used to establish repetitive generations and for isolation of vegetable oils.

The invention now being generally described, it will be more readily understood by reference to the following examples which are included for purposes of illustration only and are not intended to limit the present invention.

5

EXAMPLES

Example 1 Cuphea KAS Factor A and B Gene Cloning

Total RNA isolated from developing seeds of *Cuphea*
10 *hookeriana* and *Cuphea pullcherrima* was used for cDNA
synthesis in commercial l-based cloning vectors. For
cloning each type of KAS gene, approximately 400,000-500,000
unamplified recombinant phage were plated and the plaques
transferred to nitrocellulose. For KAS factor B cloning
15 from *C. hookeriana*, a mixed probe containing *Brassica napus*
KAS factor B and *Ricinus communis* (Castor) KAS factor B
radiolabeled cDNA's was used. Similarly, a mixed probe
containing *Brassica napus* KAS factor A and *Ricinus communis*
KAS factor A cDNA clones was used to obtain *C. hookeriana*
20 KAS factor A genes. For KASIII, a spinach KASIII cDNA
clone obtained from Dr. Jan Jaworski was radiolabeled and
used as a probe to isolate a KASIII clone from *C.*
hookeriana. For KAS B and KAS A cloning from *C.*
pullcherrima, *C. hookeriana* KAS B and KAS A genes chKAS B-2
25 and chKAS A-2-7 (see below) were radiolabeled and used as
probes.

DNA sequence and translated amino acid sequence for
Cuphea KAS clones are provided in Figures 1-9. *Cuphea*
hookeriana KAS factor B clones chKAS B-2 and chKAS B-31-7

are provided in Figures 1 and 2. Neither of the clones is full length. *Cuphea hookeriana* KAS Factor A clones chKAS A-2-7 and chKAS A-1-6 are provided in Figures 3 and 4. chKAS A-2-7 contains the entire encoding sequence for the KAS factor protein. Based on comparison with other plant synthase proteins, the transit peptide is believed to be represented in the amino acids encoded by nucleotides 125-466. chKAS A-1-6 is not a full length clone although some transit peptide encoding sequence is present. Nucleotides 1-180 represent transit peptide encoding sequence, and the mature protein encoding sequence is believed to begin at nucleotide 181.

Cuphea pullcherrima KAS factor B clones cpuKAS B/7-8 and cpuKAS B/8-7A are provided in Figures 5 and 6. Both of the clones contain the entire encoding sequences for the KAS factor B proteins. The first 35 amino acids of cpuKAS B/7-8 are believed to represent the transit peptide, with the mature protein encoding sequence beginning at nucleotide 233. The first 39 amino acids of cpuKAS B/8-7A are believed to represent the transit peptide, with the mature protein encoding sequence beginning at nucleotide 209. *Cuphea pullcherrima* KAS factor A clones cpuKAS A/p7-6A and cpuKAS A-p8-9A are provided in Figures 7 and 8. Both of the clones contain the entire encoding sequences for the KAS factor A proteins. Translated amino acid sequence of cpuKAS A/p7-6A is provided. The mature protein is believed to begin at the lysine residue encoded 595-597, and the first 126 amino acids are believed to represent the transit peptide. The DNA sequence of KAS A clone cpuKAS A-p8-9A is preliminary.

Further analysis will be conducted to determine final DNA sequence and reveal the amino acid sequence encoded by this gene.

DNA and translated amino acid sequence of *Cuphea*

5 *hookeriana* KASIII clone chKASIII-27 is provided in Figure 9. The encoding sequence from nucleotides 37-144 of chKASIII-27 are believed to encode a transit peptide, and the presumed mature protein encoding sequence is from nucleotides 145-1233.

10 Deduced amino acid sequence of the *C. hookeriana* KAS factor B and KAS factor A cDNA's reveals strong homology to the *Brassica napus* and *Ricinus communis* clones previously reported. The *C. hookeriana* KAS factor B clone is more homologous to the *Ricinus* and *Brassica* KAS factor B clones
15 (94% and 91% respectively) than it is to the *Ricinus* and *Brassica* KAS factor A clones (60% for both). Furthermore, the *C. hookeriana* KAS factor A clone is more homologous to the *Ricinus* and *Brassica* KAS factor A clones (85% and 82% respectively) than it is the *Ricinus* and *Brassica* KAS factor
20 B clone (60% for both). The *C. hookeriana* KAS factor B cDNAs designated as chKAS B-2 and chKAS B-31-7 are 96% identical within the mature portion of the polypeptide. Similarly, the deduced amino acid sequence of the mature protein regions of the *C. hookeriana* KAS factor A clones
25 chKAS A-2-7 and chKAS A-1-6 are 96% identical. The *C. pullcherrima* KAS clones also demonstrate homology to the *R. communis* and *Brassica napus* KAS clones. The mature protein portion of all of the KAS factor A family members in the different *Cuphea* species are 95% identical. Similarly the

mature protein portion of the KAS factor B genes in *Cuphea* are also 95-97% identical with each other. However there is only approximately 60% sequence identity between KAS factor B and KAS factor A clones either within the same or
5 different species of *Cuphea*.

Example 2 Levels and Patterns of Expression

To examine tissue specificity of KAS expression in *Cuphea hookeriana*, Northern blot analysis was conducted
10 using total RNA isolated from seed, root, leaf and flower tissue. Two separate but identical blots were hybridized with either chKAS B-31-7 or chKAS A-2-7 coding region probes. The data from this RNA blot analysis indicate that KAS B is expressed at a similar level in all tissues
15 examined, whereas KAS A expression is detected only in the seed. These results also demonstrate a different level of expression for each of the synthases. KAS A is an abundant message, whereas KAS B is expressed at low levels. Furthermore, even under highly stringent hybridization
20 conditions (65°C, 0.1 X SSC, 0.5% SDS), the KAS A probe hybridizes equally well with two seed transcripts of 2.3 and 1.9 kb. The larger hybridizing band is likely the transcript of the KAS A-2-7 gene since the size of its cDNA is 2046bp, and the number of clones obtained from cDNA
25 screening corresponds well with the apparent mobility of the mRNA and its abundance on the blot.

Example 3 Expression of Plant KAS Genes in E.coli

DNA fragments encoding the mature polypeptide of the *Cuphea hookeriana* KAS A cDNAs and the *Cuphea pullcherrima* KAS B cDNAs were obtained by PCR and cloned into a QIAexpress expression vector (Qiagene). Experimental conditions for maximum level of expression were determined for all of these clones and the parameters for highest level of soluble fraction were identified. Cells are grown in ECLB media containing 1M sorbitol and 2.5 mM betaine overnight and subcultured as a 1:4 dilution in the same medium. Cells are then grown for 2 hours (to approximately .6-.8 O.D.) and induced with 0.4 mM IPTG and allowed to grow for 5 more hours.

Enzyme activity of the affinity purified recombinant enzymes obtained from over-expression of the chKAS A-2-7 and cpuKAS B/8-7A clones was measured using a wide range of acyl-ACP substrates (6:0- to 16:1-ACP). The activity profile for cpuKAS B/8-7A is provided in Fig.10. The data demonstrate that the enzyme is active with all acyl-ACP substrates examined, although activity on 6:0 to 14:0-ACP substrates is substantially greater than the activity on 16:0 and 16:1 substrates.

The activity profile of the *C. hookeriana* KAS A clones chKAS A-2-7 and chKAS A-1-6 is provided in Figure 11. The *C. hookeriana* KAS A clones are most active with C:6, and have the least activity with C:16:0 substrates. However, the activity of this clone on even the preferred C6:0 substrate

is 50 fold lower than the activity of the *C. pullcherrima* KAS B clones.

A fragment containing the mature protein encoding portion of a *R. communis* KAS factor A clone was also cloned into a QIAexpress expression vector, expressed in *E. coli* and the enzyme affinity purified as described above. The activity profile for castor KAS A is provided in Figure 12. Highest activity is observed with C14:0 substrates, although some activity is also seen with C6:0 and C16:1. In comparison, the activity profile obtained from purified *R. communis* KAS factor B also using the QIAexpress expression system is provided in Figure 13. The KAS B clone demonstrates substantially higher levels of activity (10 fold and higher) than the *R. communis* KAS A clone. The preference of the KAS factor B for 6:0- to 14:0-ACP substrates is consistent with the previous observations that this protein provides KAS I activity.

Example 4 KAS and TE Expression in Transgenic Seed

Both the CpFatB1 (*C. hookeriana* thioesterase cDNA; Dehesh et al. (1996) *Plant Physiol.* 110:203-210) and the chKAS A-2-7 were PCR amplified, sequenced, and cloned into a napin expression cassette. The napin/cp FatB1 and the napin/KAS A-2-7 fusions were ligated separately into the binary vector pCGN1558 (McBride and Summerfelt (*Pl.Mol.Biol.* (1990) 14:269-276) and transformed into *A. tumefaciens*, EHA101. The resulting CpFatB1 binary construct is pCGN5400 and the chKAS A-2-7 construct is pCGN5401. *Agrobacterium* mediated transformation of a *Brassica napus* canola variety

was carried out as described by Radke et al. (*Theor. Appl. Genet.* (1988) 75:685-694; *Plant Cell Reports* (1992) 11:499-505). Several transgenic events were produced for each of the pCGN5400 and pCGN5401 constructs.

5 A double gene construct containing a napin/cpFatB1 expression construct in combination with a napin/chKAS A-2-7 expression construct was also assembled, ligated into a binary vector and used for co-cultivation of a canola *Brassica* variety. The binary construct containing the
10 chFatB1 and chKAS A-2-7 expression constructs is pCGN5413.

Fatty acid analysis of 26 transgenic lines containing chKAS A-2-7 (5401 lines) showed no significant changes in the oil content or profile as compared to similar analyses of wild type canola seeds of the transformed variety.

15 Fatty acid analysis of 36 transgenic lines containing cpFatB1 (5400 lines) showed increased levels of C:8 and C:10 in transgenic seeds. The highest level of C:8 observed in a pool seed sample was 4.2 mol%. The C:10 levels were between 30 and 35% of the C:8 content. Fatty acid analysis of 25
20 transgenic lines containing the TE/KAS A tandem (5413 lines) demonstrated an overall increase in both C:8 and C:10 levels relative to those observed with TE containing lines (5400) alone. In lines containing the cpFatB1 construct alone, the average level of C:8 average were 1.5 mol%, whereas the C:8
25 average levels in TE/KAS A tandem containing lines was 2.37 mol%. The ratio of C:8 to C:10 remained constant in both populations. The number of transgenic events relative to the C:8 content are presented in Figure 14. These data show that the transgenic events with tandem TE/KAS A construct

yield more lines with higher levels of C:8 than those events with single TE construct. For example, several lines containing nearly 7 mole% C8 were obtained with the TE/KAS A pCGN5413 construct, whereas the highest C8 containing line
5 from the pCGN5400 TE alone transformation contained 4.2 mole% C8.

Half seed analysis of the T3 generation of transgenic canola plants expressing a ChFatB2 (*C. hookeriana* thioesterase; Dehesh et al. (1996) *The Plant Journal* 9:167-
10 172) indicate that these plant can accumulate up to 22 weight% (33 mol%) of 8:0 and 10:0 fatty acids (4804-22-357). Segregation analysis shows that these transformants contain two loci and that they are now homozygous. Selected plants grown from these half seeds were transferred into the
15 greenhouse and later crossed with T1 transformants that had been transformed with either *Cuphea hookeriana* KAS A (5401) alone or KAS A/CpFatB1 double constructs (5413).

Fatty acid analysis of several events resulting from the crosses between transgenic lines containing ChFatB2
20 (4804-22-357) and chKAS A-2-7 (5401-9), reveal an increase in the ratio of C:10/C:8 levels (Figure 15). This C:10/C:8 ratio in nearly all of the transgenic events containing ChFatB2 TE alone fluctuates between 3 and 6, whereas in the F1 generation of transgenic containing both the TE and the
25 KAS A-2-7, the ratio can be as high as 22. This increase in C:10 levels is accompanied by an increase in the total C:8 and C:10 content (Figure 16). The sum of the C:8 and C:10 fatty acids in the heterozygous F1 lines is as high as those in the homozygous parent line (4804-22-357), whereas the

heterozygous lines usually contain substantially less C:8 and C:10 than the homozygous lines.

Similar results were observed in F1 generation seeds resulting from crosses performed between 4804-22-357

5 (ChFatB2) and the 5413-17 event (CpFatB1 and chKAS A-2-7 tandem). Levels of C:8 and C:10 in the 5413-17 line were 6.3 and 2.8 mol% respectively. Data presented in Figure 17 show that there is shift towards C:10 fatty acids as was observed with the 4804-22-357 (ChFatB2) x 5401-9 (chKAS A-2-7) crosses. Furthermore, Figure 18 indicates the presence of two separate populations of heterozygotes. Those containing approximately 9-11 weight percent C:10 + C:8 are believed to represent offspring containing a single copy of the ChFatB1 TE gene and no copies of the CpFatB1 and chKAS A genes from 5413. Those plants containing approximately 15-20 weight percent C:10 + C:8 are believed to represent the heterozygotes containing a single ChFatB1 TE gene as well as the CpFatB1 and chKAS A genes from 5413. Thus, the level of the C:10 + C:8 fatty acids does not decrease to 50% of that detected in parent lines when a copy of the ChKAS A gene is present.

To further characterize the chain length specificity of the *Cuphea hookeriana* KAS A enzyme, crosses between transgenic *Brassica napus* lines containing a California Bay (Umbellularia californica) 12:0 specific thioesterase, Uc FatB1 (USPN 5,344,771) and chKAS A-2-7 (5401-9) were made. Half seed analysis of transgenic plants containing Uc fatB1 have previously indicated that these plants can accumulate up to 52 mol% C12:0 in the seed oil of homozygous dihaploid

lines (LA86DH186). Crosses between the line LA86DH186 and untransformed control *Brassica* demonstrated a decrease in the C12:0 levels.

However, crosses between LA86DH186 and the 5401-9 hemizygous line led to an accumulation of up to 57 mol% C12:0 in the seed oil of F1 progeny (Figure 19).

Interestingly, in crosses with LA86DH186 x untransformed control line and LA86DH186 x 5401-9, levels of C14:0 in the seeds of the F1 progeny decreased to 50% of the levels

obtained in homozygous LA86DH186 lines (Figure 20).

Furthermore, increases in the proportion of C12:0 fatty acid resulted in a substantial decline in the proportions of all the long-chain fatty acyl groups (C16:0, C18:0, C18:2, and C18:3). These results indicate that the ChKAS A-2-7 is an enzyme with substrate specificity ranging from C6:0 to C10:0-ACP, and that its over-expression ultimately reduces the longer chain acyl-ACP pools.

Further evidence is obtained in support of the chain length specificity of the ChKAS A-2-7 in crosses of the

5401-9 line with a transgenic line (5266) expressing an 18:1/18:0 TE from *Garcinia mangostana* (GarmFatA1, US patent application No. 08/440,845). Transgenic *Brassica* line 5266 has been shown to accumulate up to 24 mol% C18:0 in the seed oil of homozygous lines (Figure 21). However, in the seed oil of F1 progeny of crosses between 5266 and 5401-9 levels of C18:0 were reduced to approximately 12 mol%.

Furthermore, levels of C16:0 generated from these crosses was similar to the levels obtained from the seed oil of nontransgenic control plants.

Example 5 In vitro Analysis of Plant KAS Enzymes

Seed extracts were prepared from developing seeds of nontransgenic controls or transgenic *Brassica* expressing chKAS A-2-7 as described in Slabaugh et al. (*Plant Journal*, 1998 in press) and Leonard et al. (*Plant Journal*, 1998, in press). *In vitro* fatty acid synthesis assays were performed as described by Post-Beittenmiller (*J. Biol. Chem.* (1991), 266:1858-1865). Extracts were concentrated by ammonium sulfate precipitation and desalting using P-6 columns (Bio-Rad, Hercules, CA). Reactions (65 μ l) contained 0.1M Tris/HCl (pH 8.0), 1 mM dithiothreitol, 25 mM recombinant spinach ACP1, 1 mM NADH, 2 mM NADPH, 50 μ M malonyl-CoA, 10 μ M [1-¹⁴C]acetyl-CoA (50 mCi/mmol), 1mg/ml BSA, and 0.25 mg/ml seed protein. Selected seed extracts were preincubated with cerulenin at 23°C for 10 min. Reaction products were separated on an 18% acrlamide gel containing 2.25M urea, electroblotted onto to nitrocellulose and quntitated by phosporimaging using Image QuaNT software (Molecular Dynamics, Sunnyvale, CA). Authentic acyl-ACPs were run in parallel, immunoblotted and finally detected by anti-ACP serum to confirm fatty acid chain lengths.

The results (Figure 22) indicate that the fatty acid synthesis capabilities of transgenic *Brasica* (5401-9) seed extracts was greater than that obtained from in the nontransgenic controls as measured by the relative abundance of C8:0- and C10:0-ACP at all time points tested. In addition, pretreatment of the extracts with cerulenin, markedly reduced the synthesis of longer chain fatty acids in both the transgenic and nontransgenic control seed

extracts. However, the extension of the spinach-ACP was much less inhibited in the seed extracts from the transgenic lines than in the seed extracts of nontransgenic control *Brassica*.

5 These data further support that Ch KAS A-2-7 is a condensing enzyme active on medium chain acyl-ACPs, and that expression of this enzyme in plants results in enlarged substrate pools to be hydrolyzed by medium-chain specific thioesterases. Furthermore, these data suggest that chKAS
10 A-2-7 also is a cerulenin-resistant condensing enzyme.

 All publications and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains.
15 All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

20 Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claim.

MISSING UPON TIME OF PUBLICATION

13. The construct of Claim 5 wherein said encoding sequence is cpuKAS A/p8-9A.

14. The construct of Claim 5 wherein said encoding sequence is chKASIII-27.

5 15. An improved method for producing medium-chain fatty acids in transgenic plant seeds by expression of a plant medium-chain thioesterase protein heterologous to said transgenic plant,

the improvement comprising expression of a plant synthase
10 factor protein heterologous to said transgenic plant in conjunction with expression of said plant medium-chain thioesterase, whereby the percentage of medium-chain fatty acids produced in seeds expressing both a plant synthase factor protein and a plant medium-chain thioesterase protein is
15 increased as compared to the percentage of medium-chain fatty acids produced in seeds expressing only said plant medium-chain thioesterase protein.

16. The method of Claim 15 wherein said medium-chain thioesterase protein is a ChFatB2 protein.

20 17. The method of Claim 15 wherein said medium-chain thioesterase protein is a CpFatB1 protein.

18. The method of Claim 15 wherein said medium-chain thioesterase protein is a C12 preferring thioesterase from California bay.

25 19. The method of Claim 15 wherein said plant synthase factor protein is expressed from a construct according to Claim 1.

20. The method of Claim 19 wherein said synthase factor A protein is from a *Cuphea* species.

21. The method of Claim 20 wherein said *Cuphea* species is *C. hookeriana* or *C. pullcherrima*.

22. A method of altering the medium-chain fatty acid composition in plant seeds expressing a heterologous plant medium-chain preferring thioesterase, wherein said method comprises

providing for expression of a plant synthase factor protein heterologous to said transgenic plant in conjunction with expression of a plant medium-chain thioesterase protein heterologous to said transgenic plant, whereby the composition of medium-chain fatty acids produced in said seeds is modified as compared to the composition of medium-chain fatty acids produced in seeds expressing said plant medium-chain thioesterase protein in the absence of expression of said plant synthase factor protein.

23. The method of Claim 22 wherein said medium-chain thioesterase protein is a ChFatB2 protein.

24. The method of Claim 22 wherein said medium-chain thioesterase protein is a CpFatB1 protein.

25. The method of Claim 22 wherein said medium-chain thioesterase protein is a C12 preferring thioesterase from California bay.

26. The method of Claim 22 wherein said plant synthase factor protein is expressed from a construct according to Claim 1.

27. The method of Claim 26 wherein said synthase factor A protein is from a *Cuphea* species.

28. The method of Claim 27 wherein said *Cuphea* species is *C. hookeriana* or *C. pullcherrima*.

29. The method of Claim 22 wherein said fatty acid composition is enriched for C10 fatty acids.

30. The method of Claim 22 wherein said fatty acid composition is enriched for C12 fatty acids.

5 31. The method of Claim 22 wherein said fatty acid composition is enriched for at least one medium chain fatty acid and at least one other medium chain fatty acid is decreased.

32. The method of Claim 31 wherein said enriched fatty
10 acid is C12 and said decreased fatty acid is C14.

y66

48 AGC TCC ACC GCG GTG GCG GCC GCT CTA GAA CTA GTG GAT CCC CCG GGC
 Ser Ser Thr Ala Val Ala Ala Leu Glu Leu Val Asp Pro Pro Gly

96 TGC AGG AAT TCG GCA CGA GCC GAT CTC GGT GCC GAC CGC CTC TCC AAG
 Cys Arg Asn Ser Ala Arg Ala Asp Leu Gly Ala Asp Arg Leu Ser Lys

144 ATC GAC AAG GAG AGA GCC GGA GTG CTG GTC GGA ACA GGA ATG GGT GGT
 Ile Asp Lys Glu Arg Ala Gly Val Leu Val Gly Thr Gly Met Gly Gly

192 CTG ACT GTC TTC TCT GAC GGG GTT CAG TCT CTT ATC GAG AAG GGT CAC
 Leu Thr Val Phe Ser Ser Asp Gly Val Gln Ser Leu Ile Glu Lys Gly His

240 CGG AAA ATC ACC CCT TTC TTC ATC CCC TAT GCC ATT ACA AAC ATG GGG
 Arg Lys Ile Thr Pro Pro Phe Phe Ile Pro Tyr Ala Ile Thr Asn Met Gly

288 TCT GCC CTG CTC GCT ATC GAA TTT GGT CTC ATG GGC CCA AAC TAT TCA
 Ser Ala Leu Leu Ala Ile Glu Phe Gly Leu Met Gly Pro Asn Tyr Ser

336 ATT TCC ACT GCA TGT GCC ACT TCC AAC TAC TGC TTC CAT GCT GCC GCT
 Ile Ser Thr Ala Cys Ala Thr Ser Ser Asn Tyr Cys Phe His Ala Ala Ala

384 AAT CAT ATC CGC CGT GGT GAG GCT GAT CTT ATG ATT GCT GGA GGC ACT
 Asn His Ile Arg Arg Gly Glu Ala Ala Asp Leu Met Ile Ala Gly Gly Thr

 FIGURE 1
 1 OF 4

2/66

432 GAG GCC GCA ATC ATT CCA ATT GGG TTG GGA GGC TTT GTG GCT TGC AGG
 Glu Ala Ala Ile Ile Pro Ile Gly Leu Gly Gly Phe Val Ala Cys Arg

480 GCT TTG TCT CAA AGG AAC GAT GAC CCG CAG ACT GCC TCT AGG CCC TGG
 Ala Leu Ser Gln Arg Asn Asp Asp Pro Gln Thr Ala Ser Arg Pro Trp

528 GAT AAA GAC CGT GAT GGT TTT GTG ATG GGT GAA GGT GCT GGA GTG TTG
 Asp Lys Asp Arg Asp Gly Phe Val Met Gly Glu Gly Ala Gly Val Leu

576 GTG ATG GAG AGC TTG GAA CAT GCA ATG AGA CGA GGA GCA CCG ATT ATT
 Val Met Glu Ser Leu Glu His Ala Met Arg Arg Gly Ala Pro Ile Ile

624 GCA GAG TAT TTG GGA GGT GCA ATC AAC TGT GAT GCT TAT CAC ATG ACT
 Ala Glu Tyr Leu Gly Gly Ala Ile Asn Cys Asp Ala Tyr His Met Thr

672 GAT CCA AGG GCT GAT GGT CTT GGT GTC TCT TCT TGC ATT GAG AGT AGC
 Asp Pro Arg Ala Asp Gly Leu Gly Val Ser Ser Cys Ile Glu Ser Ser

720 CTT GAA GAT GCT GGC GTC TCA CCT GAA GAG GTC AAT TAC ATA AAT GCT
 Leu Glu Asp Ala Gly Val Ser Pro Glu Glu Val Asn Tyr Ile Asn Ala

768 CAT GCG ACT TCT ACT CTA GCT GGG GAT CTC GCC GAG ATA AAT GCC ATC
 His Ala Thr Ser Thr Leu Ala Gly Asp Leu Ala Glu Ile Asn Ala Ile

FIGURE 1
2 OF 4

3/66

AAG AAG GTT TTC AAG AAC ACA AAG GAT ATC AAA ATT AAT GCA ACT AAG Lys Lys Val Phe Lys Asn Thr Lys Asp Ile Lys Ile Asn Ala Thr Lys	816
TCA ATG ATC GGA CAC TGT CTT GGA GCA TCT GGA GGT CTT GAA GCT ATA Ser Met Ile Gly His Cys Leu Gly Ala Ser Gly Leu Glu Ala Ile	864
GCG ACT ATT AAG GGA ATA AAC ACC GGC TGG CTT CAT CCC AGC ATT AAT Ala Thr Ile Lys Gly Ile Asn Thr Gly Trp Leu His Pro Ser Ile Asn	912
CAA TTC AAT CCT GAG CCA TCG GTG GAG TTC GAC ACT GTT GCC AAC AAG Gln Phe Asn Pro Glu Pro Ser Val Glu Phe Asp Thr Val Ala Asn Lys	960
AAG CAG CAA CAC GAA GTT AAC GTT GCG ATC TCG AAT TCA TTC GGA TTT Lys Gln Gln His Glu Val Asn Val Ala Ile Ser Asn Ser Phe Gly Phe	1008
GGA GGC CAC AAC TCA GTC GTG GCT TTC TCG GCT TTC AAG CCA TGATTA Gly Gly His Asn Ser Val Val Ala Phe Ser Ala Phe Lys Pro	1056
CCCATTTCAC AAGGTACTTG TCATTGAGAA TACGGATTAT GGACTTGCAG AGTAATTTC	1116
CCATGTTTGT CGGAAGAGCA TATTACCAG GTTGTCGTC AAACCCATT AGGATACTGT	1176

FIGURE 1
3 OF 4

4/66

TCTATGTAAT AAAACTAAGG ATTATTAATT TCCCTTTTAA TCCTGTCTCC AGTTTGAGCA	1236
TGAAATTATA TTTATTTTAT CTTAGAAAGG TCAAATAAGA TTTTGTTTTT CCTCTGTAAA	1296
ACTTTTGTTT GTATTGGAAA GGAAGTGCCG TCTCAAAAAA AAAAAAAAAA AA	1348

FIGURE 1
4 OF 4

5/66

Sequence Range: 1 to 1704

	10	20	30	40
AAA TTA ACC CTC ACT AAA GGG AAC AAA AGC TGG AGC TCC ACC GNG GTG				
Lys Leu Thr Leu Thr Lys Gly Asn Lys Ser Trp Ser Ser Thr Xxx Val>				
50	60	70	80	90
	*			
CGG GCC GCT CTA GAA CTA GTG GAT CCC CCG GGC TGC AGG AAT TCG GCA				
Ala Ala Ala Leu Glu Leu Val Asp Pro Pro Gly Cys Arg Asn Ser Ala>				
100	110	120	130	140
		*		
CGA GCC GGC ATG GGC CTC CTC TCC GTA TTC GGC TCC GAC GTC GAC TCT				
Arg Ala Gly Met Gly Leu Val Ser Val Phe Gly Ser Asp Val Asp Ser>				
150	160	170	180	190
			*	
TAT TAC GAA AAG CTC CTC TCC GGC GAG AGC GGG ATC AGC TTA ATC GAC				
Tyr Tyr Glu Lys Leu Leu Ser Gly Glu Ser Gly Ile Ser Leu Ile Asp>				
200	210	220	230	240
				*
CGC TTC GAC GCT TCC AAG TTC CCC ACC AGG TTC GGC GGC CAG ATC CGG				
Arg Phe Asp Ala Ser Lys Phe Pro Thr Arg Phe Gly Gly Gln Ile Arg>				
250	260	270	280	
GGA TTC AAC GCG ACG GGA TAC ATC GAC GGC AAG AAC GAC AGG AGG CTC				
Gly Phe Asn Ala Thr Gly Tyr Ile Asp Gly Lys Asn Asp Arg Arg Leu>				
90	300	310	320	330
	*			
GAC GAT TGC CTC CGC TAC TGC ATT GTC GCC GGC AAG AAG GCT CTC GAA				
Asp Asp Cys Leu Arg Tyr Cys Ile Val Ala Gly Lys Lys Ala Leu Glu>				

FIGURE 2
1/5

6/66

340 350 360 370 380
 AAT TCC GAT CTC GGC GGT GAA AGC CTC TCC AAG ATT GAT AAG GAG AGA
 Asn Ser Asp Leu Gly Gly Glu Ser Leu Ser Lys Ile Asp Lys Glu Arg>

 390 400 410 420 430
 GCT GGA GTG CTA GTT GGA ACT GGT ATG GGT GGC CTA ACC GTC TTC TCT
 Ala Gly Val Leu Val Gly Thr Gly Met Gly Gly Leu Thr Val Phe Ser>

 440 450 460 470 480
 GAC GGG GTT CAG AAT CTC ATC GAG AAA GGT CAC CGG AAG ATC TCC CCG
 Asp Gly Val Gln Asn Leu Ile Glu Lys Gly His Arg Lys Ile Ser Pro>

 490 500 510 520
 TTT TTC ATT CCC TAT GCC ATT ACA AAC ATG GGG TCT GCT CTG CTT GCC
 Phe Phe Ile Pro Tyr Ala Ile Thr Asn Met Gly Ser Ala Leu Leu Ala>

 530 540 550 560 570
 ATC GAT TTG GGT CTG ATG GGC CCA AAC TAT TCG ATT TCA ACT GCA TGT
 Ile Asp Leu Gly Leu Met Gly Pro Asn Tyr Ser Ile Ser Thr Ala Cys>

 580 590 600 610 620
 GCT ACT TCC AAC TAC TGC TTT TAT GCC GCT GCC AAT CAT ATC CGC CGA
 Ala Thr Ser Asn Tyr Cys Phe Tyr Ala Ala Ala Asn His Ile Arg Arg>

 630 640 650 660 670
 GGC GAG GCT GAC CTC ATG ATT GCT GGA GGA ACT GAG GCT GCA ATC ATT
 Gly Glu Ala Ala Asp Leu Met Ile Ala Gly Gly Thr Glu Ala Ala Ile Ile>

FIGURE 2
2/5

7/66

680	690	700	710	720
CCA ATT GGG TTA GGA GGA TTC GTT GCC TGC AGG GCT TTA TCT CAA AGG				*
Pro Ile Gly Leu Gly Gly Phe Val Ala Cys Arg Ala Leu Ser Gln Arg>				
730	740	750	760	
AAT GAT GAC CCT CAG ACT GCC TCA AGG CCG TGG GAT AAG GAC CGT GAT				
Asn Asp Asp Pro Gln Thr Ala Ser Arg Pro Trp Asp Lys Asp Arg Asp>				
70	780	790	800	810
	*			
GGT TTT GTG ATG GGC GAA GGG GCT GGA GTA TTG GTT ATG GAG AGC TTG				
Gly Phe Val Met Gly Glu Gly Ala Gly Val Leu Val Met Glu Ser Leu>				
820	830	840	850	860
		*		
GAA CAT GCA ATG AAA CGA GGA GCG CCG ATT ATT GCA GAA TAT TTG GGA				
Glu His Ala Met Lys Arg Gly Ala Pro Ile Ile Ala Glu Tyr Leu Gly>				
870	880	890	900	910
			*	
GGT GCA GTC AAT TGT GAT GCT TAT CAT ATG ACT GAT CCA AGG GCT GAT				
Gly Ala Val Asn Cys Asp Ala Tyr His Met Thr Asp Pro Arg Ala Asp>				
920	930	940	950	960
				*
GGG CTT GGT GTC TCC TCT TGC ATT GAG AGC AGT CTG GAA GAT GCT GGG				
Gly Leu Gly Val Ser Ser Cys Ile Glu Ser Ser Leu Glu Asp Ala Gly>				
970	980	990	1000	
GTC TCA CCT GAA GAG GTC AAT TAC ATA AAT GCT CAT GCG ACT TCC ACT				
Val Ser Pro Glu Glu Val Asn Tyr Ile Asn Ala His Ala Thr Ser Thr>				

FIGURE 2
3/5

8/66

10	1020	1030	1040	1050
	CTT GCT GGG GAT CTT GCC GAG ATA AAT GCC ATC AAG AAG GGT TTC AAG			
	Leu Ala Gly Asp Leu Ala Glu Ile Asn Ala Ile Lys Lys Val Phe Lys>			
1060	1070	1080	1090	1100
	AAC ACC AAG GAA ATC ACA ATC AAT GCA ACT AAG TCG ATG ATC GGA CAC			
	Asn Thr Lys Glu Ile Thr Ile Asn Ala Thr Lys Ser Met Ile Gly His>			
1110	1120	1130	1140	1150
	TGT CTT GGA GCA TCA GGG GGT CTT GAA GCC ATT GCG ACA ATT AAG GGA			
	Cys Leu Gly Ala Ser Gly Gly Leu Glu Ala Ile Ala Thr Ile Lys Gly>			
1160	1170	1180	1190	1200
	ATA ACC ACC GGC TGG CTT CAT CCC AGC ATA AAC CAA TTC AAT CCC GAG			
	Ile Thr Thr Gly Trp Leu His Pro Ser Ile Asn Gln Phe Asn Pro Glu>			
	1210	1220	1230	1240
	CCA TCA GTG GAA TTC GAC ACA GTT GCC AAC AAG AAG CAG CAA CAT GAA			
	Pro Ser Val Glu Phe Asp Thr Val Ala Asn Lys Lys Gln Gln His Glu>			
50	1260	1270	1280	1290
	GTG AAT GTT GCT ATC TCA AAT TCA TTC GGA TTC GGA GGC CAC AAC TCA			
	Val Asn Val Ala Ile Ser Asn Ser Phe Gly Phe Gly Gly His Asn Ser>			
1300	1310	1320	1330	1340
	GTT GTA GCT TTC TCA GCC TTC AAG CCA TGA TTA CTC GGT TCA AAT GCA			
	Val Val Ala Phe Ser Ala Phe Lys Pro			

FIGURE 2
4/5

9/66

AATTGTGTC TGAGACAGTG AGCTTCAACT TGCAGAGCAA TTTTITACAT GCCTTGTCGT
CGGAAGAGCG TAATACCGGG ATAGTTCCTT GATAGTTCAT TTAGGATGTT TTAGTGCAAT
AATCGAAGAT TATTTCATT CTAATCCAGT CTCCGNCGAG TTGAGAATC TATCTGTTTG
TATTAGAAAG AACGAGGCAA GATTTTGTTT CATGTTGTG TTGTATTAC TTTCTTTTG
CCCTTGTCAG TGGCATTTAA GATAAGCTTA TAAAAAATAA AAAAAATAA AAAACTCGAG
GGGGGGCCCG GTACCCAATT CGCCCTATAG TGAGTCGTAT GACAATTCAC TGTCCTCGG

FIGURE 2
5/5

10/66

10 20 30 40 50 60
 ACTAAAGGGA ACAAAGCTG GAGCTCCACC GCGGTGGCGG CCGCTCTAGA ACTAGTGGAT *

 70 80 90 100 110 120
 CCCCCGGGCT GCAGGAATTC GGCACGAGTT TTCTTACTTG GGTCGGCTCA GCTCAGGTGT *

 130 140 150 160
 TCCA ATG GCG ACC GCT TCT TGC ATG GTT GCG TCC CCT TTC TGT ACG TGG
 Met Ala Thr Ala Ser Cys Met Val Ala Ser Pro Phe Cys Thr Trp

 170 180 190 200 210
 CTC GTA GCT GCA TGC ATG CCC ACT TCA TCC GAC AAC GAC CCA CGT TCC
 Leu Val Ala Ala Cys Met Pro Thr Ser Ser Asp Asn Asp Pro Arg Ser

 220 230 240 250 260
 CTT TCC CAC AAG CGG CTC CGC CTC TCC CGT CGC CGG AGG ACT CTC TCC
 Leu Ser His Lys Arg Leu Arg Leu Ser Arg Arg Arg Thr Leu Ser

 270 280 290 300 310
 TCC CAT TGC TCC CTC CGC GGA TCC ACC TTC CAA TGC CTC GAT CCT TGC
 Ser His Cys Ser Leu Arg Gly Ser Thr Phe Gln Cys Leu Asp Pro Cys

 320 330 340 350 360
 AAC CAG CAA CGC TTC CTC GGG GAT AAC GGA TTC GCT TCC CTC TTC GGA *
 Asn Gln Gln Arg Phe Leu Gly Asp Asn Gly Phe Ala Ser Leu Phe Gly

11/66

370 380 390 400
 TCC AAG CCT CTT CGT TCA AAT CGC GGC CAC CTG AGG CTC GGC CGC ACT
 Ser Lys Pro Leu Arg Ser Asn Arg Gly His Leu Arg Leu Gly Arg Thr

 410 420 430 440 450
 TCC CAT TCC GGG GAG GTC ATG GCT GTG GCT ATG CAA CCT GCA CAG GAA
 Ser His Ser Gly Glu Val Met Ala Val Ala Met Gln Pro Ala Gln Glu

 460 470 480 490 500
 GTC TCC ACA AAT AAG AAA CCT GCT ACC AAG CAA AGG CGA GTA GTT GTG
 Val Ser Thr Asn Lys Lys Pro Ala Thr Lys Gln Arg Arg Val Val Val

 510 520 530 540 550
 ACA GGT ATG GGC GTG GTG ACT CCT CTA GGC CAT GAC CCC GAT GTT TAC
 Thr Gly Met Gly Val Val Thr Pro Leu Gly His Asp Pro Asp Val Tyr

 560 570 580 590 600
 TAC AAC AAT CTC CTA GAC GGA ATA AGT GGC ATA AGT GAG ATA GAG AAC
 Tyr Asn Asn Leu Leu Asp Gly Ile Ser Gly Ile Ser Glu Ile Glu Asn

 610 620 630 640
 TTC GAC TGC TCT CAG TTT CCC ACG AGA ATT GCC GGA GAG ATC AAG TCT
 Phe Asp Cys Ser Ser Gln Phe Pro Thr Arg Ile Ala Gly Glu Ile Lys Ser

 650 660 670 680 690
 TTT TCC ACA GAT GGC TGG GTG GCC CCA AAG TTC TCC GAG AGG ATG GAC
 Phe Ser Thr Asp Gly Trp Val Ala Pro Lys Phe Ser Glu Arg Met Asp

FIGURE 3 2 OF 6

12/66

700	710	720	730	740
AAG TTC ATG CTT TAC ATG CTG ACT GCA GGC AAG AAA GCA TTA GCA GAT				
Lys Phe Met Leu Tyr Met Leu Thr Ala Gly Lys Lys Ala Leu Ala Asp				
750	760	770	780	790
GGT GGA ATC ACT GAA GAT GCG ATG AAA GAG CTC AAT AAA AGA AAG TGT			*	
Gly Gly Ile Thr Glu Asp Ala Met Lys Glu Leu Asn Lys Arg Lys Cys				
800	810	820	830	840
GGA GTT CTC ATT GGC TCC GGA TTG GGC GGT ATG AAG GTA TTC AGC GAT				*
Gly Val Leu Ile Gly Ser Gly Leu Gly Met Lys Val Phe Ser Asp				
850	860	870	880	
TCC ATT GAA GCT CTG AGG ACT TCA TAT AAG AAG ATC AGT CCC TTT TGT				
Ser Ile Glu Ala Leu Arg Thr Ser Tyr Lys Lys Ile Ser Pro Phe Cys				
890	900	910	920	930
GTA CCT TTT TCT ACC ACA AAT ATG GGA TCC GCT ATT CTT GCA ATG GAC				
Val Pro Phe Ser Thr Thr Asn Met Gly Ser Ala Ile Leu Ala Met Asp				
940	950	960	970	980
TTG GGA TGG ATG GGC CCT AAC TAT TCG ATA TCA ACT GCC TGT GCA ACA		*		
Leu Gly Trp Met Gly Pro Asn Tyr Ser Ile Ser Thr Ala Cys Ala Thr				
990	1000	1010	1020	1030
AGT AAC TTC TGT ATA CTG AAT GCT GCG AAC CAC ATA ATC AAA GGC GAA			*	
Ser Asn Phe Cys Ile Leu Asn Ala Ala Asn His Ile Ile Lys Gly Glu				

FIGURE 3 3 OF 6

13/66

1040	1050	1060	1070	1080-
<p>* GCA GAC ATG ATG CTT TGT GGT GGC TCG GAT GCG GCC GTT TTA CCT GTT Ala Asp Met Met Leu Cys Gly Ser Asp Ala Ala Val Leu Pro Val</p>				
1090	1100	1110	1120	
<p>GGT TTG GGA GGT TTC GTA GCA TGC CGA GCT TTG TCA CAG AGG AAT AAT Gly Leu Gly Gly Phe Val Ala Cys Arg Ala Leu Ser Gln Arg Asn Asn</p>				
1130	1140	1150	1160	1170
<p>* GAC CCT ACC AAA GCT TCG AGA CCA TGG GAC AGT AAT CGT GAT GGA TTT Asp Pro Thr Lys Ala Ser Arg Pro Trp Asp Ser Asn Arg Asp Gly Phe</p>				
1180	1190	1200	1210	1220
<p>* GTG ATG GGA GAA GGA GCT GGA GTT TTA CTT CTT GAG GAG TTA GAG CAT Val Met Gly Gly Ala Gly Val Leu Leu Leu Glu Glu Leu Glu His</p>				
1230	1240	1250	1260	1270
<p>* GCA AAG AAA AGA GGT GCA ACC ATT TAT GCG GAA TTT CTA GGT GCG AGT Ala Lys Lys Arg Gly Ala Thr Ile Tyr Ala Glu Phe Leu Gly Gly Ser</p>				
1280	1290	1300	1310	1320
<p>* TTC ACT TGC GAC GCC TAC CAC ATG ACC GAG CCT CAC CCT GAA GGA GCT Phe Thr Cys Asp Ala Tyr His Met Thr Glu Pro His Pro Glu Gly Ala</p>				
1330	1340	1350	1360	
<p>GGT GTG ATC CTC TGC ATA GAG AAG GCC TTG GCT CAG TCC GGA GTC TCG Gly Val Ile Leu Cys Ile Glu Lys Ala Leu Ala Gln Ser Gly Val Ser</p>				

FIGURE 3 4 OF 6

14/66

1370 1380 1390 1400 1410
 *
 AGG GAA GAC GTA AAT TAC ATA AAT GCG CAT GCA ACT TCC ACT CCT GCT
 Arg Glu Asp Val Asn Tyr Ile Asn Ala His Ala Thr Ser Thr Pro Ala

 1420 1430 1440 1450 1460
 *
 GGA GAT ATC AAG GAA TAC CAA GCT CTC GCC CAC TGT TTC GGC CAA AAC
 Gly Asp Ile Lys Glu Tyr Gln Ala Leu Ala His Cys Phe Gly Gln Asn

 1470 1480 1490 1500 1510
 *
 AGT GAG CTG AGA GTG AAT TCC ACC AAA TCG ATG ATC GGT CAC CTT CTT
 Ser Glu Leu Arg Val Asn Ser Thr Lys Ser Met Ile Gly His Leu Leu

 1520 1530 1540 1550 1560
 *
 GGA GGA GCT GGT GGC GTA GAA GCA GTT GCA GTA GTT CAG GCA ATA AGG
 Gly Gly Ala Gly Gly Val Glu Ala Val Ala Val Val Gln Ala Ile Arg

 1570 1580 1590 1600
 ACA GGA TGG ATC CAT CCA AAT ATT AAT TTG GAA GAC CCG GAC GAA GGC
 Thr Gly Trp Ile His Pro Asn Ile Asn Leu Glu Asp Pro Asp Glu Gly

 1610 1620 1630 1640 1650
 *
 GTG GAT GCA AAA CTG CTC GTC GGC CCT AAG AAG GAG AAA CTG AAG GTC
 Val Asp Ala Lys Leu Leu Val Gly Pro Lys Lys Glu Lys Leu Lys Val

 1660 1670 1680 1690 1700
 *
 AAG GTC GGT TTG TCC AAT TCA TTT GGG TTC GGC GGC CAT AAC TCA TCC
 Lys Val Gly Leu Ser Asn Ser Phe Gly Phe Gly Gly His Asn Ser Ser

FIGURE 3 5 OF 6

15/66

1710	1720	1730	1740	1750	1760
ATA CTA TTT GCC CCC TGC AAC TAG A AAAGAGTCTG TGGAAGCCGA GAGTCTTTGA			*		
Ile Leu Phe Ala Pro Cys Asn ***					
1770	1780	1790	1800	1810	1820
GAACTCATGC ACGTTAGTAG CTTCTTATGC CTCTGAAACC GAGATAGACC GGCTACTCGA			*		
1830	1840	1850	1860	1870	1880
GGGGATGCCA AAGATACTCC TTGCCGGTAT TGGTGTTAAG AGATCACTGC TTGTCCCTTT			*		
1890	1900	1910	1920	1930	1940
TATTTTCTTC TTCTTTTGAG AGCTTTAACC GAGGTAGTCG TATTTTCGAG CTTTTCGAAT			*		
1950	1960	1970	1980	1990	2000
ACATGTTTCGT TATCGGATCA ATGTGTTTCT TCTAAGATCA TTTGTAATGC ATATTTTGAA			*		
2010	2020	2030	2040		
AAACCCACATC TCAGTATGCA AAATAAAAAA AAAAAAAAAA AAAAAA			*		

FIGURE 3 6 OF 6

16/66

Sequence Range: 1 to 1921

10	20	30	40	50	60
CGGCACGAGG	TCACCTCTTA	CCTCGCCTGC	TTCGAGCCCT	GCCATGACTA	CTACACCTCC
70	80	90	100	110	120
GCATCCTTGT	TCGGATCCAG	GCCCATCCGC	ACCACCCGCA	GGCACCGGAG	GCTCAATCGA
130	140	150	160	170	180
GCTTCCCTTT	CCGGGGAGGC	AATGGCTGTG	GCTCTGCAAC	CTGCACAGGA	AGTTACCACA
190	200	210	220		
AAG AAG AAG	CCA AGT ATC	AAA CAG CGG	CGA GTA GTT	GTG ACT GGA	ATG
Lys Lys Lys	Pro Ser Ile	Lys Gln Arg	Arg Val Val	Val Thr Gly	Met>
230	240	250	260	270	
GGT GTG GTG	ACT CCT CTA	GGC CAT GAC	CCT GAT GTT	TTC TAC AAT	AAT
Gly Val Val	Thr Pro Leu	Gly His Asp	Pro Asp Val	Phe Tyr Asn	Asn>
280	290	300	310	320	
CTG CTT GAT	GGA ACG AGT	GGC ATA AGT	GAG ATA GAG	ACC TTT GAT	TGT
Leu Leu Asp	Gly Thr Ser	Gly Ile Ser	Glu Ile Glu	Thr Phe Asp	Cys>
330	340	350	360	370	
GCT CAA TTT	CCT ACG AGA	ATT GCT GGA	GAG ATC AAG	TCT TTC TCC	ACA
Ala Gln Phe	Pro Thr Arg	Ile Ala Gly	Glu Ile Lys	Ser Phe Ser	Thr>

FIGURE 4
1/6

17/66

380	390	400	410	420
GAT GGT TGG GTG GCC CCG AAG CTC TCC AAG AGG ATG GAC AAG TTC ATG				*
Asp Gly Trp Val Ala Pro Lys Leu Ser Lys Arg Met Asp Lys Phe Met>				
430	440	450	460	
CTT TAC ATG CTG ACT GCC GGC AAG AAA GCA TTA ACA AAT GGT GGA ATC				
Leu Tyr Met Leu Thr Ala Gly Lys Lys Ala Leu Thr Asn Gly Gly Ile>				
470	480	490	500	510
	*			
ACC GAA GAT GTG ATG AAA GAG CTA GAT AAA AGA AAA TGC GGA GTT CTC				
Thr Glu Asp Val Met Lys Lys Glu Leu Asp Lys Arg Lys Cys Gly Val Leu>				
520	530	540	550	560
		*		
ATT GGC TCA GCA ATG GGT GGA ATG AAG GTA TTC AAT GAT GCC ATT GAA				
Ile Gly Ser Ala Met Gly Gly Met Lys Val Phe Asn Asp Ala Ile Glu>				
570	580	590	600	610
			*	
GCC CTA AGG ATT TCA TAT AAG AAG ATG AAT CCC TTT TGT GTA CCT TTC				
Ala Leu Arg Ile Ser Tyr Lys Lys Met Asn Pro Phe Cys Val Pro Phe>				
620	630	640	650	660
				*
GCT ACC ACA AAT ATG GGA TCA GCT ATG CTT GCA ATG GAC TTG GGA TGG				
Ala Thr Thr Asn Met Gly Ser Ala Met Leu Ala Met Asp Leu Gly Trp>				
670	680	690	700	
ATG GGC CCC AAC TAC TCG ATA TCT ACT GCT TGT GCA ACG AGT AAC TTT				
Met Gly Pro Asn Tyr Ser Ile Ser Thr Ala Cys Ala Thr Ser Asn Phe>				

FIGURE 4
2/6

18 166

710 720 730 740 750
 TGT ATC CTG AAT GCT GCG AAC CAC ATA ATC AGA GGC GAA GCA GAT GTG
 Cys Ile Leu Asn Ala Ala Asn His Ile Ile Arg Gly Glu Ala Asp Val>
 760 770 780 790 800
 ATG CTT TGC GGG GGC TCA GAT GCG GTA ATC ATA CCT ATT GGT ATG GGA
 Met Leu Cys Gly Gly Ser Asp Ala Val Ile Ile Pro Ile Gly Met Gly>
 810 820 830 840 850
 GGT TTT GTT GCA TGC CGA GCT TTG TCA CAG AGA AAT GCC GAC CCT ACT
 Gly Phe Val Ala Cys Arg Ala Leu Ser Gln Arg Asn Ala Asp Pro Thr>
 860 870 880 890 900
 AAA GCT TCA AGA CCA TGG GAC AGT AAT CGT GAT GGA TTT GTT ATG GGG
 Lys Ala Ser Arg Pro Trp Asp Ser Asn Arg Asp Gly Phe Val Met Gly>
 910 920 930 940
 GAA GGA GCT GGA GTG CTA CTA CTA GAG GAG TTA GAG CAT GCA AAG AAA
 Glu Gly Ala Gly Val Leu Leu Leu Glu Glu Leu Glu His Ala Lys Lys>
 950 960 970 980 990
 AGA GGT GCG ACT ATT TAC GCA GAA TTT CTA GGT GGA AGT TTC ACT TGC
 Arg Gly Ala Thr Ile Tyr Ala Glu Phe Leu Gly Gly Ser Phe Thr Cys>
 1000 1010 1020 1030 1040
 GAT GCC TAC CAC ATG ACC GAG CCT CAC CCT GAT GGA GCT GGA GTG ATT
 Asp Ala Tyr His Met Thr Glu Pro His Pro Asp Gly Ala Gly Val Ile>

FIGURE 4

3/6

19 / 26

1050	1060	1070	1080	1090
CTC TGC ATA GAG AAG GCT TTG GCT CAG TCA GGA GTC TCT AGG GAA GAC			*	
Leu Cys Ile Glu Lys Ala Leu Ala Gln Ser Gly Val Ser Arg Glu Asp>				
1100	1110	1120	1130	1140
GTA AAT TAC ATA AAT GCA CAT GCC ACA TCC ACT CCA GCT GGA GAT ATC				*
Val Asn Tyr Ile Asn Ala His Ala Thr Ser Thr Pro Ala Gly Asp Ile>				
1150	1160	1170	1180	
AAA GAG TAC CAA GCT CTT ATC CAC TGT TTC GGC CAA AAC AAC GAG TTA				
Lys Glu Tyr Gln Ala Leu Ile His Cys Phe Gly Gln Asn Asn Glu Leu>				
1190	1200	1210	1220	1230
	*			
AAA GTG AAT TCT ACC AAA TCA ATG ATT GGT CAC CTT CTC GGA GCA GCC				
Lys Val Asn Ser Thr Lys Ser Met Ile Gly His Leu Leu Gly Ala Ala>				
1240	1250	1260	1270	1280
		*		
GGT GGT GTG GAA GCA GTT TCA GTA GTT CAG GCA ATA AGG ACT GGG TGG				
Gly Gly Val Glu Ala Val Ser Val Val Gln Ala Ile Arg Thr Gly Trp>				
1290	1300	1310	1320	1330
			*	
ATC CAT CCG AAT ATT AAT TTG GAA AAC CCA GAT GAA GGC GTG GAT ACC				
Ile His Pro Asn Ile Asn Leu Glu Asn Pro Asp Glu Gly Val Asp Thr>				
1340	1350	1360	1370	1380
				*
AAA TTG CTC GTG GGC CCT AAG AAG GAG AGA CTG AAC ATT AAG GTC GGT				
Lys Leu Leu Val Gly Pro Lys Lys Glu Arg Leu Asn Ile Lys Val Gly>				

FIGURE 4
4/6

20/66

1390 1400 1410 1420
 TTG TCT AAT TCA TTC GGG TTT GGT GGG CAC AAC TCG TCC ATA CTC TTC
 Leu Ser Asn Ser Phe Gly Phe Gly Gly His Asn Ser Ser Ile Leu Phe>
 1430 1440 1450 1460 1470 1480
 GCC CCT TAC AAC TAG GCGGTTT CATGTGTGGA ATTCTACTCA ATCTATCAAA
 Ala Pro Tyr Asn ***>
 1490 1500 1510 1520 1530 1540
 GCTGAAGTTT TGAGGACTCC AGCATGTTGG TAGCTCCTTA CGTCTCTAGA CATGCCCATG
 1550 1560 1570 1580 1590 1600
 AGTTTGTGT CGGGAGCTGT AGTCGGAACC ATGACGGATT GAGTACTCAT GCGCACACAG
 1610 1620 1630 1640 1650 1660
 GATATACTCC TTGCTAGAAAT TGTTAGAGCA CTATTCATTA TCCCATTTTT TTTCTGAAAT
 1670 1680 1690 1700 1710 1720
 CTCCCTCCTT ACGGTAGTTG TACTTTCGAG CGTTTCATCG AGTCAGTGAA GAAGAGAACA
 1730 1740 1750 1760 1770 1780
 AAGCTAACTC GGCACGCTAG TAACCATTTG CCCTTTGTTT TGCTCTCTAT TTTATCGCCG
 1790 1800 1810 1820 1830 1840
 TTTTGTGGGT TAAAATTGT AAAACTAGAC GACTGGTTTG TTTTCTCTTG ATCATTTGGAG

FIGURE 4
5/6

21 / 66

1850	1860	1870	1880	1890	1900
	*				
ATGTATGGCC	ATATTGCCT	TTCATTGATG	ATAAAAAAAAA	AAAAAAAAAAA	AAAAAAAAAAA
1910	1920				
	*				
AAAAAAAAAA	AAAAAAAAAA	A			

FIGURE 4
6/6

22/66

CTGGTACGCC TGCAGGTACC GGTCCGGAAT TCCCGGGTCG ACCACGCGT CCGTCTTCCC 60
 ACTCCGATCG TTCTTCTTCC ACCGCATCTC TTCTCTTCTC TTGGCTTCTC CGCCATCCTC 120
 CGCCGCC ATG CAT TCC CTC CAG TCA CCC TCC CTT CGG GCC TCC CCG CTC 169
 Met His Ser Leu Gln Ser Pro Ser Leu Arg Ala Ser Pro Leu 10
 1
 GAC CCC TTC CGC CCC AAA TCA TCC ACC GTC CGC CCC CTC CAC CGA GCA 217
 Asp Pro Phe Arg Pro Lys Ser Ser Thr Val Arg Pro Leu His Arg Ala 30
 15 20 25 30
 TCA ATT CCC AAC GTC CGG GCC GCT TCC CCC ACC GTC TCC GCT CCC AAG 265
 Ser Ile Pro Asn Val Arg Ala Ala Ser Pro Thr Val Ser Ala Pro Lys 40 45
 35
 CGC GAG ACC GAC CCC AAG AAG CGC GTC GTG ATC ACC GGA ATG GGC CTT 313
 Arg Glu Thr Asp Pro Lys Lys Arg Val Val Ile Thr Gly Met Gly Leu 50 55 60
 GTC TCC GTT TTC GGC TCC GAC GTC GAT GCG TAC TAC GAC AAG CTC CTG 361
 Val Ser Val Phe Gly Ser Asp Val Asp Ala Tyr Tyr Asp Lys Leu Leu 65 70 75
 TCA GGC GAG AGC GGG ATC GGC CCA ATC GAC CGC TTC GAC GCC TCC AAG 409
 Ser Gly Glu Ser Gly Ile Gly Pro Ile Asp Arg Phe Asp Ala Ser Lys 80 85 90
 TTC CCC ACC AGG TTC GGC GGC CAG ATT CGT GGC TTC AAC TCC ATG GGA 457
 Phe Pro Thr Arg Phe Gly Gly Gln Ile Arg Gly Phe Asn Ser Met Gly 95 100 105 110
 TAC ATT GAC GGC AAA AAC GAC AGG CGG CTT GAT GAT TGC CTT CGC TAC 505
 Tyr Ile Asp Gly Lys Asn Asp Arg Arg Leu Asp Cys Leu Arg Tyr 115 120 125

FIGURE 5
1/4

23/66

TGC ATT GTC GCC GGG AAG AAG TCT CTT GAG GAC GCC GAT CTC GGT GCC 553
 Cys Ile Val Ala Gly Lys Lys Ser Leu Glu Asp Ala Asp Leu Gly Ala
 130 135 140
 GAC CGC CTC TCC AAG ATC GAC AAG GAG AGA GCC GGA GTG CTG GTT GGG 601
 Asp Arg Leu Ser Lys Ile Asp Lys Glu Arg Ala Gly Val Leu Val Gly
 145 150 155
 ACA GGA ATG GGT GGT CTG ACT GTC TTC TCT GAC GGG GTT CAA TCT CTT 649
 Thr Gly Met Gly Gly Leu Thr Val Phe Ser Asp Gly Val Gln Ser Leu
 160 165 170
 ATC GAG AAG GGT CAC CGG AAA ATC ACC CCT TTC TTC ATC CCC TAT GCC 697
 Ile Glu Lys Gly His Arg Lys Ile Thr Pro Phe Phe Ile Pro Tyr Ala
 175 180 185 190
 ATT ACA AAC ATG GGG TCT GCC CTG CTC GCT ATT GAA CTC GGT CTG ATG 745
 Ile Thr Asn Met Gly Ser Ala Leu Leu Ala Ile Glu Leu Gly Leu Met
 195 200 205
 GGC CCA AAC TAT TCA ATT TCC ACT GCA TGT GCC ACT TCC AAC TAC TGC 793
 Gly Pro Asn Tyr Ser Ile Ser Thr Ala Cys Ala Thr Ser Asn Tyr Cys
 210 215 220
 TTC CAT GCT GCT AAT CAT ATC CGC CGT GGT GAG GCT GAT CTT ATG 841
 Phe His Ala Ala Ala Asn His Ile Arg Arg Gly Glu Ala Asp Leu Met
 225 230 235
 ATT GCT GGA GGC ACT GAG GCC GCA ATC ATT CCA ATT GGG TTG GGA GGC 889
 Ile Ala Gly Gly Thr Glu Ala Ala Ile Ile Pro Ile Gly Leu Gly Gly
 240 245 250

FIGURE 5
2/4

24/66

TTT GTG GCT TGC AGG GCT CTG TCT CAA AGG AAC GAT GAC CCT CAG ACT 937
 Phe Val Ala Cys Arg Ala Leu Ser Gln Arg Asn Asp Asp Pro Gln Thr 270
 255 260 265
 GCC TCT AGG CCC TGG GAT AAA GAC CGT GAT GGT TTT GTG ATG GGT GAA 985
 Ala Ser Arg Pro Trp Asp Lys Asp Arg Asp Gly Phe Val Met Gly Glu 285
 275 280
 GGT GCT GGA GTG TTG GTG CTG GAG AGC TTG GAA CAT GCA ATG AAA CGA 1033
 Gly Ala Gly Val Leu Val Leu Glu Ser Leu Glu His Ala Met Lys Arg 300
 290 295
 GGA GCA CCT ATT ATT GCA GAG TAT TTG GGA GGT GCA ATC AAC TGT GAT 1081
 Gly Ala Pro Ile Ile Ala Glu Tyr Leu Gly Gly Ala Ile Asn Cys Asp 315
 305 310
 GCT TAT CAC ATG ACT GAC CCA AGG GCT GAT GGT CTC GGT GTC TCC TCT 1129
 Ala Tyr His Met Thr Asp Pro Arg Ala Asp Gly Leu Gly Val Ser Ser 330
 320 325
 TGC ATT GAG AGT AGC CTT GAA GAT GCT GGC GTC TCA CCT GAA GAG GTC 1176
 Cys Ile Glu Ser Ser Leu Glu Asp Ala Gly Val Ser Pro Glu Glu Val 350
 335 340 345
 AAT TAC ATA AAT GCT CAT GCG ACT TCT ACT CTA GCT GGG GAT CTC GCC 1224
 Asn Tyr Ile Asn Ala His Ala Thr Ser Thr Leu Ala Gly Asp Leu Ala 365
 355 360
 GAG ATA AAT GCC ATC AAG AAG GTT TTC AAG AAC ACA AAG GAT ATC AAA 1272
 Glu Ile Asn Ala Ile Lys Lys Val Phe Lys Asn Thr Lys Asp Ile Lys 380
 370 375

FIGURE 5
3/4

25/166

```

ATT AAT GCA ACT AAG TCA ATG ATC GGA CAC TGT CTT GGA GCC TCT GGA      1320
Ile Asn Ala Thr Lys Ser Met Ile Gly His Cys Leu Gly Ala Ser Gly
385 390 395

GGT CTT GAA GCT ATA GCG ACT ATT AAG GGA ATA AAC ACC GGC TGG CTT      1368
Gly Leu Glu Ala Ile Ala Thr Ile Lys Gly Ile Asn Thr Gly Trp Leu
400 405 410

CAT CCC AGC ATT AAT CAA TTC AAT CCT GAG CCA TCC GTG GAG TTC GAC      1416
His Pro Ser Ile Asn Gln Phe Asn Pro Glu Pro Ser Val Glu Phe Asp
415 420 425 430

ACT GTT GCC AAC AAG AAG CAG CAA CAC GAA GTT AAT GTT GCG ATC TCG      1464
Thr Val Ala Asn Lys Lys Gln Gln His Glu Val Asn Val Ala Ile Ser
435 440 445

AAT TCA TTT GGA TTC GGA GGC CAC AAC TCA GTC GTG GCT TTC TCG GCT      1512
Asn Ser Phe Gly Phe Gly Gly His Asn Ser Val Val Ala Phe Ser Ala
450 455 460

TTC AAG CCA TGA TTACC CATTTCACAA GGCACCTGTC ATTGAGAGTA CGGTGTGTCG      1569
Phe Lys Pro

TCAAACCCAT TTAGGATACT GTTCTATGTA AAAAAAAGTA AGGATTATCA CTTCCCTTC      1629

TAATCCTGTC TCCAGTTTGA GAATGAAATT ATATTATTTT TAAAAAAGGC      1689

GGCCGCTCTA GAGGATCCAA GCT      1712

```

FIGURE 5
4/4

26 / 66

Sequence Range: 1 to 1802

10	20	30	40	50	60
GGTCGACCCA	CGCGTCCGGG	CTTCCGACC	ACATTTCATT	TCTTGCCCTCG	TTATCTCCGC
70	80	90	100	110	
CGCTCCTCCG	CGTCGTTCG	CGCGCGCCG	C ATG CAA	TCC CTC CAC	TCC CCT TCC
			Met Gln	Ser Leu	His Ser Pro Ser
120	130	140	150	160	
CTC CGC CCC	TCC CCT CTC	GAG CCC TTC	CGC CTC AAT	TCC CCC TCC	TCC
Leu Arg Pro	Ser Pro Leu	Glu Pro Phe	Arg Leu Asn	Ser Pro Ser	Ser
170	180	190	200	210	
GCC GCC GCT	CTC CGC CCC	CGT CGC GCC	AGC CTC CCC	GTC ATC	CGT
Ala Ala Ala	Leu Arg Pro	Leu Arg Ala	Ser Leu Pro	Val Ile	Arg
220	230	240	250		
GCT GCC ACC	GCC TCC GCC	CCC AAG	CGC GAG TCC	GAC CCC	AAG AAG CGG
Ala Ala Thr	Ala Ser Ala	Pro Lys	Arg Glu Ser	Asp Pro	Lys Lys Arg
260	270	280	290	300	
GTC GTC ATC	ACC GGC ATG	GGC CTC GTC	TCC GTC TTC	GGC TCC	GAC GTC
Val Val Ile	Thr Gly Met	Gly Leu Val	Ser Val Phe	Gly Ser	Asp Val
310	320	330	340	350	
GAC GCC TAC	TAC GAC AAG	CTG CTC TCC	GGC GAG AGC	GGC ATC	AGC CTA
Asp Ala Tyr	Tyr Asp Lys	Leu Leu Ser	Gly Glu Ser	Gly Ile	Ser Leu

FIGURE 6
1/5

27/66

360	370	380	390	400
*				
ATC GAC CGC TTC GAC GCT TCC AAA TTC CCC ACC AGG TTC GCC GGC CAG				
Ile Asp Arg Phe Asp Ala Ser Lys Phe Pro Thr Arg Phe Ala Gly Gln				
410	420	430	440	450
*				
ATC CGT GGC TTC AAC GCG ACG GGC TAC ATC GAC GGC AAG AAC GAC CGG				
Ile Arg Gly Phe Asn Ala Thr Gly Tyr Ile Asp Gly Lys Asn Asp Arg				
460	470	480	490	
*				
CGG CTC GAC GAT TGC CTC CGC TAC TGC ATT GTC GCC GGC AAG AAG GCT				
Arg Leu Asp Asp Cys Leu Arg Tyr Cys Ile Val Ala Gly Lys Lys Ala				
500	510	520	530	540
				*
CTC GAA GAC GCC GAT CTC GCC GGC CAA TCC CTC TCC AAG ATT GAT AAG				
Leu Glu Asp Ala Asp Leu Ala Gly Gln Ser Leu Ser Lys Ile Asp Lys				
550	560	570	580	590
GAG AGG GCC GGA GTG CTA GTT GGA ACC GGT ATG GGT GGC CTA ACT GTC				
Glu Arg Ala Gly Val Leu Val Gly Thr Gly Met Gly Gly Leu Thr Val				
600	610	620	630	640
*				
TTC TCT GAC GGG GTT CAG AAT CTC ATC GAG AAA GGT CAC CGG AAG ATC				
Phe Ser Asp Gly Val Gln Asn Leu Ile Glu Lys Gly His Arg Lys Ile				
650	660	670	680	690
*				
TCC CCG TTT TTC ATT CCA TAT GCC ATT ACA AAC ATG GGG TCT GCG CTG				
Ser Pro Phe Phe Ile Pro Tyr Ala Ile Thr Asn Met Gly Ser Ala Leu				

FIGURE 6
2/5

28/66

700	710	720	730
CTT GCC ATC GAT TTG GGT CTG ATG GGC CCA AAC TAT TCG ATT TCA ACT		*	
Leu Ala Ile Asp Leu Gly Leu Met Gly Pro Asn Tyr Ser Ile Ser Thr			
740	750	760	770
			780
GCA TGT GCT ACT TCC AAC TAC TGC TTT TAT GCT GCC GCC AAT CAT ATC			*
Ala Cys Ala Thr Ser Asn Tyr Cys Phe Tyr Ala Ala Ala Asn His Ile			
790	800	810	820
			830
CGC CGA GGT GAG GCT GAC CTG ATG ATT GCT GGA GGA ACT GAG GCT GCG			
Arg Arg Gly Glu Ala Asp Leu Met Ile Ala Gly Gly Thr Glu Ala Ala			
840	850	860	870
*			880
GTC ATT CCA ATT GGT TTA GGA GGA TTC GTT GCC TGC AGG GCT TTA TCT			
Val Ile Pro Ile Gly Leu Gly Gly Phe Val Ala Cys Arg Ala Leu Ser			
890	900	910	920
	*		930
CAA AGG AAT GAT GAT CCT CAG ACT GCC TCA AGG CCG TGG GAT AAG GAC			
Gln Arg Asn Asp Asp Pro Gln Thr Ala Ser Arg Pro Trp Asp Lys Asp			
940	950	960	970
		*	
CGT GAT GGC TTT GTG ATG GGT GAA GGG GCT GGA GTA TTG GTT ATG GAG			
Arg Asp Gly Phe Val Met Gly Glu Gly Ala Gly Val Leu Val Met Glu			
980	990	1000	1010
			1020
AGC TTG GAG CAT GCA ATG AAA CGG GGA GCG CCG ATT ATT GCA GAA TAT			*
Ser Leu Glu His Ala Met Lys Arg Gly Ala Pro Ile Ile Ala Glu Tyr			

FIGURE 6
3/5

29/66

1030 1040 1050 1060 1070
 TTG GGA GGT GCA GTC AAC TGT GAT GCT TAT CAT ATG ACT GAT CCA AGG
 Leu Gly Gly Ala Val Asn Cys Asp Ala Tyr His Met Thr Asp Pro Arg
 1080 1090 1100 1110 1120
 *
 GCT GAT GGG CTT GGT GTC TCC TCG TGC ATT GAG AGC AGT CTC GAA GAT
 Ala Asp Gly Leu Gly Val Ser Ser Cys Ile Glu Ser Ser Leu Glu Asp
 1130 1140 1150 1160 1170
 *
 GCC GGG GTC TCA CCT GAA GAG GTC AAT TAC ATA AAT GCT CAT GCG ACT
 Ala Gly Val Ser Pro Glu Glu Val Asn Tyr Ile Asn Ala His Ala Thr
 1180 1190 1200 1210
 *
 TCT ACT CTT GCT GGG GAT CTT GCC GAG ATA AAT GCC ATT AAG AAA GTT
 Ser Thr Leu Ala Gly Asp Leu Ala Glu Ile Asn Ala Ile Lys Lys Val
 1220 1230 1240 1250 1260
 *
 TTC AAG AAC ACC AAG GAA ATC AAA ATC AAT GCA ACT AAG TCA ATG ATC
 Phe Lys Asn Thr Lys Glu Ile Lys Ile Asn Ala Thr Lys Ser Met Ile
 1270 1280 1290 1300 1310
 GGA CAC TGT CTT GGA GCA TCA GGA GGT CTT GAA GCC ATC GCA ACC ATT
 Gly His Cys Leu Gly Ala Ser Gly Gly Leu Glu Ala Ile Ala Thr Ile
 1320 1330 1340 1350 1360
 *
 AAG GGA ATA ACC ACC GGC TGG CTT CAT CCC AGC ATT AAT CAA TTT AAT
 Lys Gly Ile Thr Thr Gly Trp Leu His Pro Ser Ile Asn Gln Phe Asn

FIGURE 6
4/5

30/66

1370 1380 1390 1400 1410
 CCC GAG CCA TCG GTG GAC TTC AAC ACT GTT GCC AAC AAA AAG CAG CAA
 Pro Glu Pro Ser Val Asp Phe Asn Thr Val Ala Asn Lys Lys Gln Gln

 1420 1430 1440 1450
 CAT GAA GTG AAC GTC GCT ATC TCG AAT TCT TTT GGA TTT GGA GGG CAC
 His Glu Val Asn Val Ala Ile Ser Asn Ser Phe Gly Phe Gly Gly His

 1460 1470 1480 1490 1500 1510
 AAC TCG GTT GTG GCA TTC TCA GCT TTC AAG CCA TGA ATTCT ACTTGTTCA
 Asn Ser Val Val Ala Phe Ser Ala Phe Lys Pro ***

 1520 1530 1540 1550 1560 1570
 AAATGCACAC CAGTTGCTGA GATAGGGCTT CAACCTTGCAG AGCAATTTT TAAATGCCTT

 1580 1590 1600 1610 1620 1630
 GTCGGAAGAG CGTAATACCG GAATAGGTCG GTCCTTTGAT AGTTCCTCGA AGCCATTAG

 1640 1650 1660 1670 1680 1690
 GATGATGTTT TACTGTAATA ATCGAAGATG ATTCCCATTT TAAATCTAGT CTCTGATTTA

 1700 1710 1720 1730 1740 1750
 TGTATTAGAA AGACCAATGA AAGATTTTGT GTCATGTTTG TGTGTGCAAT GTTATTTAAG

 1760 1770 1780 1790 1800
 ATAAAGCAAA AAAAAAAAAA AAGGGCGGCC GCTCTAGAGG ATCCAGCTTA CT

FIGURE 6
5/5

31 / 66

Sequence Range: 1 to 2369

```

10      20      30      40      50      60
GTACGCCTGC AGGTACCGGT CCGGAATTCC CGGGTCGACC CACGCGTCCG CATAAAAGAG *

70      80      90      100     110     120
AGAGAGAGGG ATCCATCGAA TCGGGCCACC CTCCTTTTCAT CTTGCGATTCA TTACCATACC *

130     140     150     160     170     180
ATTCCGCTGA TCCATTTTCC GCCTTTTCCG GGTCTTTTCAT CCCAAAGGGT ATCCTTTTCT

190     200     210     220     230
ATCCTATCTT CTCAAAGGGT CAGTCAGTTC CCTCCA ATG CCT GCC GCC TCT TCC
Met Pro Ala Ala Ser Ser>

240     250     260     270     280
CTG CTC GCT TCC CCT CTC TGT ACG TGG CTC CTT GCC GCC TGC ATG TCT
Leu Leu Ala Ser Pro Leu Cys Thr Trp Leu Leu Ala Ala Cys Met Ser>

290     300     310     320     330
ACC TTC CAC CCC TCC GAC CCT CTT CCG CCT TCC ATC TCC TCT CCT
Thr Ser Phe His Pro Ser Asp Pro Leu Pro Pro Ser Ile Ser Ser Pro>

340     350     360     370
CGC CGA CGC CTC TCC CGC CGC CGG ATT CTC TCC CAA TGC GCC CCA CTA
Arg Arg Arg Leu Ser Arg Arg Arg Ile Leu Ser Gln Cys Ala Pro Leu>

```

FIGURE 7
1/7

32/66

```

380      CCT TCT GCT TCC TCC GCC CTC CGC GGA TCC AGT TTC CAT ACC CTC GTC
      Pro Ser Ala Ser Ser Ala Leu Arg Gly Ser Ser Phe His Thr Leu Val>
      430      440      450      460      470
      ACC TCT TAC CTC GCC TGC TTC GAG CCC TGC CAT GAC TAC TAT ACA TCC
      Thr Ser Tyr Leu Ala Cys Phe Glu Pro Cys His Asp Tyr Tyr Thr Ser>
      480      490      500      510      520
      *
      GCA TCC TTG TTC GGA TCC AGA CCC ATT CGC ACC ACC CGC AGG CAC CGG
      Ala Ser Leu Phe Gly Ser Arg Pro Ile Arg Thr Thr Arg Arg His Arg>
      530      540      550      560      570
      AGG CTC AAT CGA GCT TCC CCT TCC AGG GAG GCA ATG GCC GTG GCT CTG
      Arg Leu Asn Arg Ala Ser Pro Ser Arg Glu Ala Met Ala Val Ala Leu>
      580      590      600      610
      CAA CCT GAA CAG GAA GTT ACC ACA AAG AAG CCA AGT ATC AAA CAG
      Gln Pro Glu Gln Glu Val Thr Thr Lys Lys Lys Pro Ser Ile Lys Gln>
      620      630      640      650      660
      CGG CGA GTA GTT GTG ACT GGA ATG GGT GTG GTG ACT CCT CTA GGC CAT
      Arg Arg Val Val Val Thr Gly Met Gly Val Val Thr Pro Leu Gly His>
      670      680      690      700      710
      GAC CCT GAT GTT TTC TAC AAT AAT CTG CTT GAT GGA ACG AGT GGC ATA
      Asp Pro Asp Val Phe Tyr Asn Asn Leu Leu Asp Gly Thr Ser Gly Ile>

```

FIGURE 7
2/7

33/66

720	730	740	750	760
* AGC GAG ATA GAG ACC TTT GAT TGT GCT CAA TTT CCT ACG AGA ATT GCT Ser Glu Ile Glu Thr Phe Asp Cys Ala Gln Phe Pro Thr Arg Ile Ala>				
770	780	790	800	810
GGA GAG ATC AAG TCT TTC TCC ACA GAT GGT TGG GTG GCC CCG AAG CTC Gly Glu Ile Lys Ser Phe Ser Thr Asp Gly Trp Val Ala Pro Lys Leu>				
820	830	840	850	
TCT AAG AGG ATG GAC AAG TTC ATG CTA TAC ATG CTG ACC GCT GGC AAG Ser Lys Arg Met Asp Lys Phe Met Leu Tyr Met Leu Thr Ala Gly Lys>		*		
860	870	880	890	900
AAA GCA TTA ACA GAT GGT GGA ATC ACC GAA GAT GTG ATG AAA GAG CTA Lys Ala Leu Thr Asp Gly Gly Ile Thr Glu Asp Val Met Lys Glu Leu>				*
910	920	930	940	950
GAT AAA AGA AAA TGC GGA GTT CTC ATT GGC TCA GCA ATG GGT GGA ATG Asp Lys Arg Lys Cys Gly Val Leu Ile Gly Ser Ala Met Gly Gly Met>				
960	970	980	990	1000
* AAG GTA TTC AAT GAT GCC ATT GAA GCC CTA AGG ATT TCA TAT AAG AAG Lys Val Phe Asn Asp Ala Ile Glu Ala Leu Arg Ile Ser Tyr Lys Lys>				
1010	1020	1030	1040	1050
ATG AAT CCC TTT TGT GTA CCT TTC GCT ACC ACA AAT ATG GGA TCA GCT Met Asn Pro Phe Cys Val Pro Phe Ala Thr Thr Asn Met Gly Ser Ala>	*			

FIGURE 7
3/7

34/66

1060	1070	1080	1090
ATG CTT GCA ATG GAC TTG GGA TGG ATG GGG CCC AAC TAC TCG ATA TCT			
Met Leu Ala Met Asp Leu Gly Trp Met Gly Pro Asn Tyr Ser Ile Ser>			
1100	1110	1120	1130
			1140 *
ACT GCT TGT GCA ACG AGT AAC TTT TGT ATA ATG AAT GCT GCG AAC CAT			
Thr Ala Cys Ala Thr Ser Asn Phe Cys Ile Met Asn Ala Ala Asn His>			
1150	1160	1170	1180
			1190
ATA ATC AGA GGC GAA GCA GAT GTG ATG CTT TGC GGG GGC TCA GAT GCG			
Ile Ile Arg Gly Glu Ala Asp Val Met Leu Cys Gly Gly Ser Asp Ala>			
1200	1210	1220	1230
			1240
GTA ATC ATA CCT ATT GGT ATG GGA GGT TTT GTT GCA TGC CGA GCT TTG			
Val Ile Ile Pro Ile Gly Met Gly Gly Phe Val Ala Cys Arg Ala Leu>			
1250	1260	1270	1280
			1290
TCC CAG AGA AAT TCC GAC CCT ACT AAA GCT TCA AGA CCA TGG GAC AGT			
Ser Gln Arg Asn Ser Asp Pro Thr Lys Ala Ser Arg Pro Trp Asp Ser>			
1300	1310	1320	1330
AAT CGT GAT GGA TTT GTT ATG GGG GAA GGA GCT GGA GTG CTA CTA CTA			
Asn Arg Asp Gly Phe Val Met Gly Glu Gly Ala Gly Val Leu Leu Leu>			
1340	1350	1360	1370
			1380 *
GAG GAG TTG GAG CAT GCA AAG AAA AGA GGT GCG ACT ATT TAC GCA GAA			
Glu Glu Leu Glu His Ala Lys Lys Arg Gly Ala Thr Ile Tyr Ala Glu>			

FIGURE 7
4/7

35/66

1390 1400 1410 1420 1430
 TTT CTA GGT GGG AGT TTC ACT TGC GAT GCC TAC CAC ATG ACC GAG CCT
 Phe Leu Gly Gly Ser Phe Thr Cys Asp Ala Tyr His Met Thr Glu Pro>
 1440 1450 1460 1470 1480
 CAC CCT GAT GGA GCT GGA GTG ATT CTC TGC ATA GAG AAG GCT TTG GCT
 His Pro Asp Gly Ala Gly Val Ile Leu Cys Ile Glu Lys Ala Leu Ala>
 1490 1500 1510 1520 1530
 CAG TCA GGA GTC TCT AGG GAA GAC GTA AAT TAC ATA AAT GCC CAT GCC
 Gln Ser Gly Val Ser Arg Glu Asp Val Asn Tyr Ile Asn Ala His Ala>
 1540 1550 1560 1570
 ACA TCC ACT CCG GCT GGA GAT ATC AAA GAG TAC CAA GCT CTT ATC CAC
 Thr Ser Thr Pro Ala Gly Asp Ile Lys Glu Tyr Gln Ala Leu Ile His>
 1580 1590 1600 1610 1620
 TGT TTC GGC CAA AAC AGA GAG TTA AAA GTT AAT TCA ACC AAA TCA ATG
 Cys Phe Gly Gln Asn Arg Glu Leu Lys Val Asn Ser Thr Lys Ser Met>
 1630 1640 1650 1660 1670
 ATT GGT CAC CTT CTC GGA GCA GCC GGT GGT GTG GAA GCA GTT TCA GTA
 Ile Gly His Leu Leu Gly Ala Ala Gly Gly Val Glu Ala Val Ser Val>
 1680 1690 1700 1710 1720
 GTT CAG GCA ATA AGG ACT GGG TGG ATC CAT CCG AAT ATT AAT TTG GAA
 Val Gln Ala Ile Arg Thr Gly Trp Ile His Pro Asn Ile Asn Leu Glu>

FIGURE 7
5/7

36166

1730 1740 1750 1760 1770
 AAC CCA GAT GAA GGC GTG GAT ACA AAA TTG CTC GTG GGT CCT AAG AAG
 Asn Pro Asp Glu Gly Val Asp Thr Lys Leu Leu Val Gly Pro Lys Lys>

 1780 1790 1800 1810
 GAG AGA CTG AAC GTT AAG GTC GGT TTG TCT AAT TCA TTT GGG TTT GGT
 Glu Arg Leu Asn Val Lys Val Gly Leu Ser Asn Ser Phe Gly Phe Gly>

 1820 1830 1840 1850 1860 1870
 GGG CAC AAC TCG TCC ATA CTC TTC GCC CCT TAC ATC TAG GAC GTTCCCGTGT
 Gly His Asn Ser Ser Ile Leu Phe Ala Pro Tyr Ile ***>

 1880 1890 1900 1910 1920 1930
 GTGGAATTCT ACTCAACATA TCAAAGCTGA AGTTTGGAG ACTCCAGCAT GTTGGTAGCT
 1940 1950 1960 1970 1980 1990
 CCTTACGCT CTAGACATGC CCATGAGTTT TGTGTCCGGA GCTTTAGTCG GAACCATGAC
 2000 2010 2020 2030 2040 2050
 GGATTGAGTA CTCATGGCGA CACTTGATAT ACTCCTTGCT AGAATTGTTG GTAGAGCAAT
 2060 2070 2080 2090 2100 2110
 ATTCAATTATC TCATATTTTT TTTTCTCTG AAATCTCCCT CCTTGAATA GTTGTACTTT
 2120 2130 2140 2150 2160 2170
 CGAGCTTTTC ATCGAGTCAG TGAAGAAGAG AACAAAGCTG TTAACCTCGG CACGTAGTAA

FIGURE 7
6/7

37 / 66

2180	2190	2200	2210	2220	2230
CCATTGCCC	TTTGTTTTC	TCTCTATTTC	ATCACCGTTT	TGTGGTTTTA	AAATTGTAA
				*	
2240	2250	2260	2270	2280	2290
AACTAGAAGA	CTGGTTTGA	TTGGTTTGT	TTCTCATTTGA	TAATTGGGGR	ATGTATGTTT
				*	
2300	2310	2320	2330	2340	2350
TGGAAATAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA
				*	
2360					
AGGGCGGCGG	CTCTAGAGG				

FIGURE 7
7/7

38 / 66

Sequence Range: 1 to 2374

10 20 30 40 50 60 *
-A-NTGGTC CGGAATTCCC GGGTCGACCC ACGCGTCCGC GAGGCCAACC CACACCAAAC
70 80 90 100 110 120 *
TTCCTCAGCT TCTCTTCTCA AGACGGAGCG CATTGGCAGC AGACAGACAG ACAGACAGAC
130 140 150 160 170 180 *
CCATAAAAGA GAGAGAGAGG GATCCATCGA ATGCGGCCAC CCTCCTTTCA TCTTCGATTC
190 200 210 220 230 240 *
ATTACCATAC CATTCGCGTG ATCCATTTC CGCCTTTTCC GGGTCTTTCA TCCCAAAGGG
250 260 270 280 290 300 *
TATCCTTTTC TATCCTATCT TCTCAAAGG TCAGTCAGTT CCCTCCAATG CCTGCCGCCT
310 320 330 340 350 360 *
CTTCCCCTGCT CGCTTCCCCCT CTCTGTACGT GGCTCCTTGC CGCCTGCATG TCTACCTCCT
370 380 390 400 410 420 *
TCCACCCCTC CGACCCCTCTT CCGCCTTCCA TCTCCTCTCC TCGCCGACGC CTCTCCCGCC
430 440 450 460 470 480 *
GCCGGATTCT CTCCCAATGC GCCCCACTAC CTTCTGCTTC CTCCGCCCTC CGCGGATCCA

FIGURE 8
1/5

39/66

490 500 510 520 530 540
 GTTTCCATAC CCTCGTCACC TCTTACCTCG CCTGCTTCGA GCCCTGCCAT GACTACTATA *
 550 560 570 580 590 600
 CATCCGCATC CTTGTTCCGA TCCAGACCCA TTCGCACCAC CCGCAGGCAC CGGAGGGCTCA *
 610 620 630 640 650 660
 ATCGAGCTTC CCCTTCCAGG GGAGGCAATG GCGGTGGCTC TGCAACCTGA ACAGGAAGTT *
 670 680 690 700 710 720
 ACCACAAAAG AGAAGCCAAG TATCAAAACAG CGGCGAGTAG TTGTGACTGG AATGGGTGTG *
 730 740 750 760 770 780
 GTGACTCCTC TAGGCCATGA ACCTGATGTT TTTCTACAAT AATCTGCTTG ATGGAACGAG *
 790 800 810 820 830 840
 TGGCATAAGC GAGATAGAGA CCTTTGATTG TGCTCAATTT CCTACGAGAA TTGCTGGAGA *
 850 860 870 880 890 900
 GATCAAGTCT TTCTCCACAG ATGGTTGGGT GGCCCCGAAG CTCTCTAAGA GGATGGACAA *
 910 920 930 940 950 960
 GTTCATGCTA TACATGCTGA CTGCTGGCAA GAAAGCATTA ACAGATGGTG GAATCACC GA *
 970 980 990 1000 1010 1020
 AGATGTGATG AAAGAGCTAG ATAAAAGAAA ATGCGGAGTT CTCATTGGCT CAGCAATGGG *

FIGURE 8
2/5

40/66

1030	1040	1050	1060	1070	1080 *
TGGAATGAAG	GTATTCAATG	ATGCCATTGA	AGCCCTAAGG	ATTTCATATA	AGAAGATGAA
1090	1100	1110	1120	1130	1140 *
TCCCTTTTGT	GTACCTTTTCG	CTACCACAAA	TATGGGATCA	GCTATGCTTG	CAATGGACTT
1150	1160	1170	1180	1190	1200 *
GGGATGGATG	GGGCCCAACT	ACTCGATATC	TACTGCTTGT	GCAACGAGTA	ACTTTTGAT
1210	1220	1230	1240	1250	1260 *
AATGAATGCT	GCGAACCATA	TAATCAGAGG	CGAAGCAGAT	GTGATGCTTT	GCGGGGGCTC
1270	1280	1290	1300	1310	1320 *
AGATGCGGTA	ATCATACCTA	TTGGTATGGG	AGGTTTGTG	GCAIGCCGAG	CTTTGTCCCA
1330	1340	1350	1360	1370	1380 *
GAGAAATCC	GACCCCTACTA	AAGCTTCAAG	ACCATGGGAC	AGTAATCGTG	ATGGATTGT
1390	1400	1410	1420	1430	1440 *
TATGGGGGAA	GGAGCTGGAG	TGCTACTACT	AGAGGAGTTG	GAGCATGCAA	AGAAAAGAGG
1450	1460	1470	1480	1490	1500 *
TGCGACTATT	TACGCAGAAT	TTCTAGGTGG	GAGTTTCACT	TGCGATGCCT	ACCACATGAC

FIGURE 8
3/5

41/66

1510	1520	1530	1540	1550	1560 *
CGAGCCTCAC	CCTGATGGAG	CTGGAGTGAT	TCTCTGCATA	GAGAAGGCTT	TGGCTCAGTC
1570	1580	1590	1600	1610	1620 *
AGGAGTCTCT	AGGGAAGACG	TAAATTACAT	AAATGCCCAT	GCCACATCCA	CTCCGGCTGG
1630	1640	1650	1660	1670	1680 *
AGATATCAAA	GAGTACCAAG	CTCTTATCCA	CTGTTTCGGC	CAAAACAGAG	AGTTAAAAAGT
1690	1700	1710	1720	1730	1740 *
TAAATTCACC	AAATCAATGA	TTGGTCACCT	TCTCGGAGCA	GCCGGTGGTG	TGGAAGCAGT
1750	1760	1770	1780	1790	1800 *
TTCAGTAGTT	CAGGCAATAA	GGACTGGGTG	GATCCATCCG	AATATTAATT	TGGAAAAACCC
1810	1820	1830	1840	1850	1860 *
AGATGAAGGC	GTGGATACAA	AATGCTCTCGT	GGGTCCCTAAG	AAGGAGAGAC	TGAACGTTAA
1870	1880	1890	1900	1910	1920 *
GGTCGGTTTG	TCTAATTTCAT	TTGGGTTTGG	TGGGCACAAC	TCGTCCATAC	TCCTCGCCCC
1930	1940	1950	1960	1970	1980 *
TTACATCTAG	GACGTTTTCGT	GTGTGGAATT	CTACTCAACA	TATCAAAGCT	GAAGTTTGA
1990	2000	2010	2020	2030	2040 *
GGACTCCAGC	ATGTTGGTAG	CTCCTTACGT	CTCTAGACAT	GCCCATGAGT	TTTGTGTCCG

FIGURE 8
4/5

42/66

2050 2060 2070 2080 2090 2100 *
GAGCTTTAGT CGGAACCATG ACGGATTGAG TACTCATGGC GACACTTGAT ATACTCCTTG
2110 2120 2130 2140 2150 2160 *
CTAGAATTGT TGGTAGAGCA ATATTCAATTA TCTCATATTT TTTTTTTCTC TGAAATCTCC
2170 2180 2190 2200 2210 2220 *
CTCCTTGCAA TAGTTGTACT TTCGAGCTTT TCATCGAGTC AGTGAAGAAG AGAACAAAGC
2230 2240 2250 2260 2270 2280 *
TGTTAACTCG GGCACGTAGT AACCATTTGC CCTTTGTTTT GCTCTCTATT TCATCACCGT
2290 2300 2310 2320 2330 2340 *
TTTGTGGTTT TAAAATTGT AAAACTAGAA GACTGGTTTA GATTGGTTTG TTTTCTCAAA
2350 2360 2370
AAAAAAAAA AAGGGCGGCC GCTCTAGAGG ATCC

FIGURE 8
5/5

43/66

Sequence Range: 1 to 1580

```

10      20      30      40      50
CCTGAATCGG ATTCAAGAGA GAGTTTCGTT GCTGGG ATG GCG AAT GCA TCT GGG
Met Ala Asn Ala Ser Gly>

60      70      80      90      100
*
TTT CTG GGT TCT TCA GTT CCT GCC CTG AGA AGG GCA ACT CAG CAT TCG
Phe Leu Gly Ser Ser Val Pro Ala Leu Arg Arg Ala Thr Gln His Ser>

110     120     130     140     150
*
ATT TCA TCG TCT CGT GGA TCT TCC TCG GAG TTT GTC TCC AAA AGG GTG
Ile Ser Ser Ser Arg Gly Ser Ser Ser Glu Phe Val Ser Lys Arg Val>

160     170     180     190
*
TTT TGC TGT AGT GCC GTT CAG GAT TCT GAC AGG CAG TCT TTG GGT GAT
Phe Cys Cys Ser Ala Val Gln Asp Ser Asp Arg Gln Ser Leu Gly Asp>

200     210     220     230     240
*
TCT CGC TCG CCG AGG CTT GTG AGT AGA GGA TGC AAA TTA ATT GGA TCT
Ser Arg Ser Pro Arg Leu Val Ser Arg Gly Cys Lys Leu Ile Gly Ser>

250     260     270     280     290
GGT TCT GCT ATA CCA GCT CTT CAA GTC TCA AAT GAT GAT CTT GCT AAA
Gly Ser Ala Ile Pro Ala Leu Gln Val Ser Asn Asp Asp Leu Ala Lys>

300     310     320     330     340
*
ATT GTC GAC ACC AAT GAT GAA TGG ATT ACT ACT GTC CGA ACG GGG ATC CGC
Ile Val Asp Thr Asn Asp Glu Trp Ile Thr Val Arg Thr Gly Ile Arg>

```

FIGURE 9
1/5

44/66

350 360 370 380 390
 AAC CGA AGG GTT CTC TCA GGT AAA GAT AGT CTT ACA AAT TTA GCA TCA
 Asn Arg Arg Val Leu Ser Gly Lys Asp Ser Leu Thr Asn Leu Ala Ser>
 400 410 420 430
 GAG GCA GCA AGG AAA GCT CTA GAG ATG GCA CAG GTA GAC GCA AAT GAT
 Glu Ala Ala Arg Lys Ala Leu Glu Met Ala Gln Val Asp Ala Asn Asp>
 440 450 460 470 480
 GTG GAT ATG GTT TTG ATG TGT ACT TCT ACC CCT GAG GAC CTT TTC GGC
 Val Asp Met Val Leu Met Cys Thr Ser Thr Pro Glu Asp Leu Phe Gly>
 490 500 510 520 530
 AGT GCT CCT CAG ATA TCG AAA GCA CTT GGC TGC AAA AAG AAT CCT TTG
 Ser Ala Pro Gln Ile Ser Lys Ala Leu Gly Cys Lys Lys Asn Pro Leu>
 540 550 560 570 580
 TCT TAC GAC ATT ACC GCT GCA TGC AGT GGA TTT GTG TTG GGT TTA GTC
 Ser Tyr Asp Ile Thr Ala Ala Cys Ser Gly Phe Val Leu Gly Leu Val>
 590 600 610 620 630
 TCA GCT GCT TGC CAC ATT AGA GGT GGG GGT TTT AAC AAT ATT CTA GTG
 Ser Ala Ala Cys His Ile Arg Gly Gly Gly Phe Asn Asn Ile Leu Val>
 640 650 660 670
 ATT GGT GCT GAT TCT CTT TCT CGG TAT GTT GAC TGG ACC GAT CGG GGA
 Ile Gly Ala Asp Ser Leu Ser Arg Tyr Val Asp Trp Thr Asp Arg Gly>

FIGURE 9

2/5

45/66

680 690 700 710 720
 ACA TGT ATT CTC TTT GGA GAT GCT GCT GGA GCT GTA GTG GTG CAG TCA
 Thr Cys Ile Leu Phe Gly Asp Ala Ala Gly Ala Val Val Gln Ser>
 730 740 750 760 770
 TGT GAT GCT GAG GAA GAT GGG CTC TTT GCT TTT GAT TTG CAT AGC GAT
 Cys Asp Ala Glu Glu Asp Gly Leu Phe Ala Phe Asp Leu His Ser Asp>
 780 790 800 810 820
 *
 GGA GAT GGG CAA AGG CAT CTA AAA GCT GCA ATC AAA GAA GAT GAA GTT
 Gly Asp Gly Gln Arg His Leu Lys Ala Ala Ile Lys Glu Asp Glu Val>
 830 840 850 860 870
 *
 GAT AAA GCC CTG GGA CAT AAT GGG TCC ATC AGA GAT TTT CCA CCA AGG
 Asp Lys Ala Leu Gly His Asn Gly Ser Ile Arg Asp Phe Pro Pro Arg>
 880 890 900 910
 *
 CGT TCT TCA TAC TCT TGC ATC CAA ATG AAC GGT AAA GAG GTA TTC CGC
 Arg Ser Ser Tyr Ser Cys Ile Gln Met Asn Gly Lys Glu Val Phe Arg>
 920 930 940 950 960
 *
 TTT GCT TGC CGC TCT GTG CCT CAG TCA ATC GAA TCA GCA CTT GGA AAG
 Phe Ala Cys Arg Ser Val Pro Gln Ser Ile Glu Ser Ala Leu Gly Lys>
 970 980 990 1000 1010
 GCC GGT CTT AAT GGA TCC AAC ATC GAC TGG TTG CTG CTT CAT CAG GCA
 Ala Gly Leu Asn Gly Ser Asn Ile Asp Trp Leu Leu His Gln Ala>

FIGURE 9

3/5

46/66

1020 * 1030 1040 1050 1060
 AAT CAG AGG ATC ATT GAT GCA GTA GCA ACA CGT CTA GAG GTT CCT CAA
 Asn Gln Arg Ile Ile Asp Ala Val Ala Thr Arg Leu Glu Val Pro Gln>
 1070 1080 1090 1100 1110
 GAA CGA ATT ATC TCA AAC TTG GCA AAT TAC GGG AAC ACT AGT GCG GCA
 Glu Arg Ile Ile Ser Asn Leu Ala Asn Tyr Gly Asn Thr Ser Ala Ala>
 1120 1130 1140 * 1150
 TCC ATT CCC TTG GCA CTA GAC GAA GCT GTG AGG AGT GGA AAT GTG AAG
 Ser Ile Pro Leu Ala Leu Asp Glu Ala Val Arg Ser Gly Asn Val Lys>
 1160 1170 1180 1190 1200 *
 CCG GGT CAC GTG ATT GCA ACC GCA GGA TTT GGC GCC GGA CTC ACA TGG
 Pro Gly His Val Ile Ala Thr Ala Gly Phe Gly Ala Gly Leu Thr Trp>
 1210 1220 1230 1240 1250 1260 *
 GGT TCT GCT ATT ATC AGG TGG GGA TAA GACTGAA GCCGAGCCAG CACTGCAGCT
 Gly Ser Ala Ile Ile Arg Trp Gly ***>
 1270 1280 1290 1300 1310 1320 *
 TCCTCTCAAA CCGATGTTTC ACGAAATTTT GCTTCCATGA CCANAAAAAG AAGAAGTCAG
 1330 1340 1350 1360 1370 1380 *
 TCTTTTATGG AGCAAGCAAC ACGACACGAT CTTTCATCACA TTGCCCTTTT TCGTTCCCTT

FIGURE 9

4/5

47/66

1390 1400 1410 1420 1430 1440 *
TTTCCATTAG TTTGATGATT TTGCTGACAA TACAATACCC ATAGTTTCTT TTGTCCCCAA
1450 1460 1470 1480 1490 1500 *
TAAGTTATTT GTTCTTGTT TAAATTGTTCA GCTTTTACTT CATTTTGTCT CGGGACATTG
1510 1520 1530 1540 1550 1560 *
GAGATGACAG CATAAACATC ATGTTTATAT TTGCTAAAA AAAAAAAAAA
1570 1580
AAAAAAAAA AAAAAAAAAA

FIGURE 9
5/5

48/26

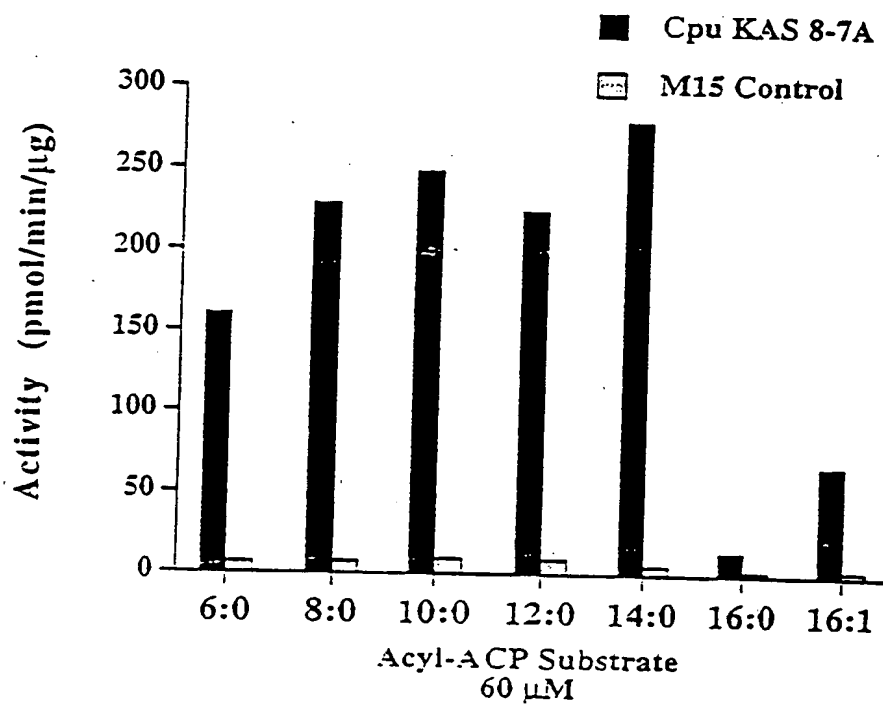


FIGURE 10

49/66

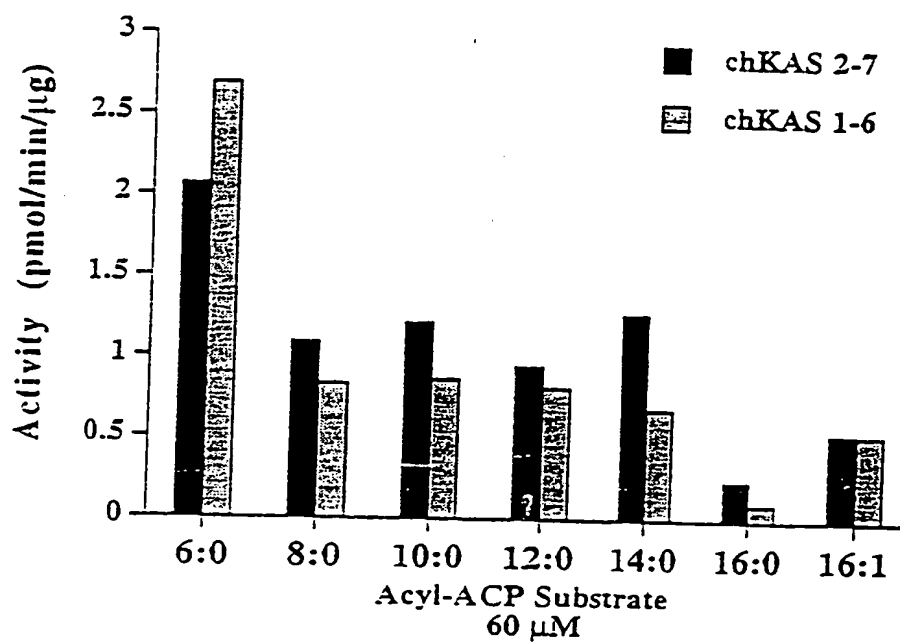


FIGURE 11

50 / 66

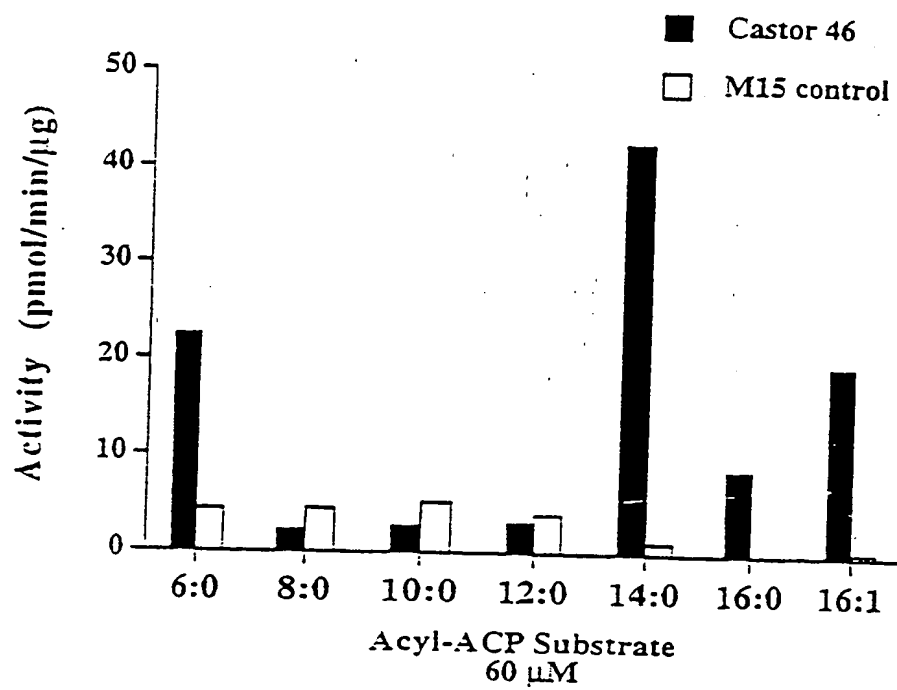
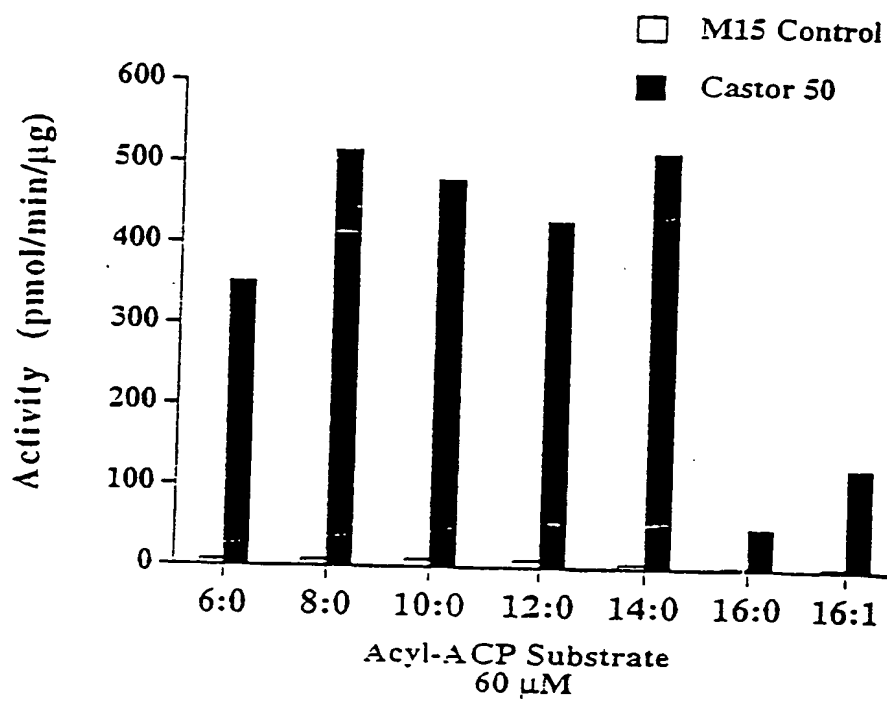


FIGURE 12

51/66



E328013-28

FIGURE 13

52/66

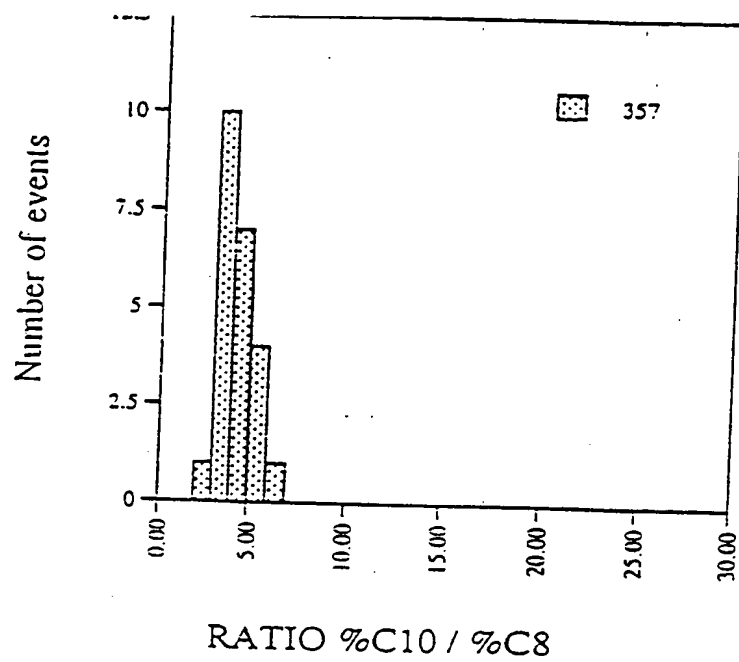


FIGURE 15

1/2

53/66

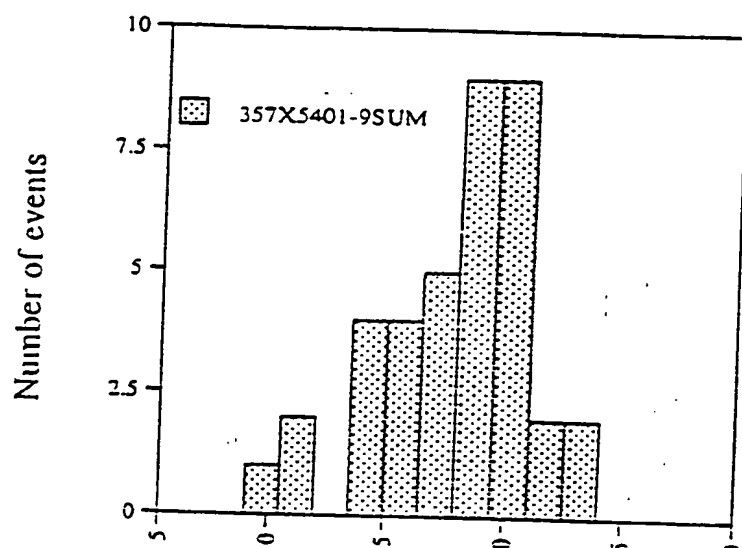


FIGURE 15
2/2

54/66

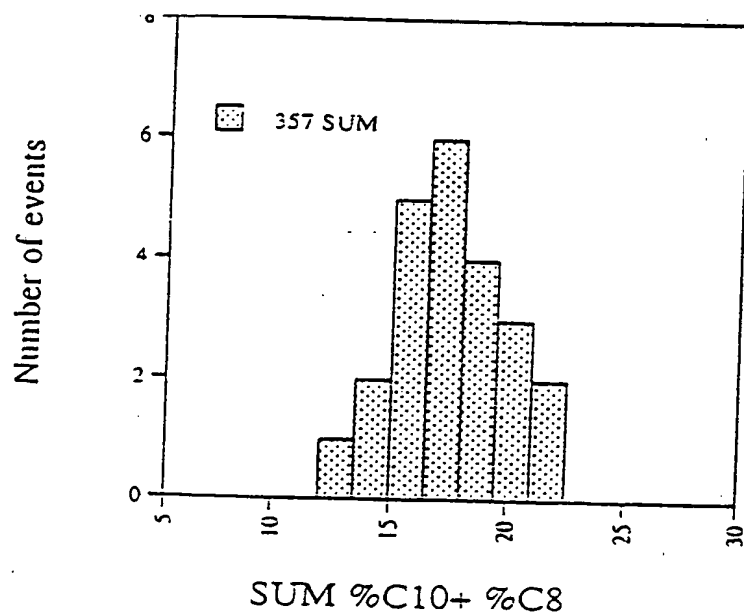


FIGURE 16

55/66

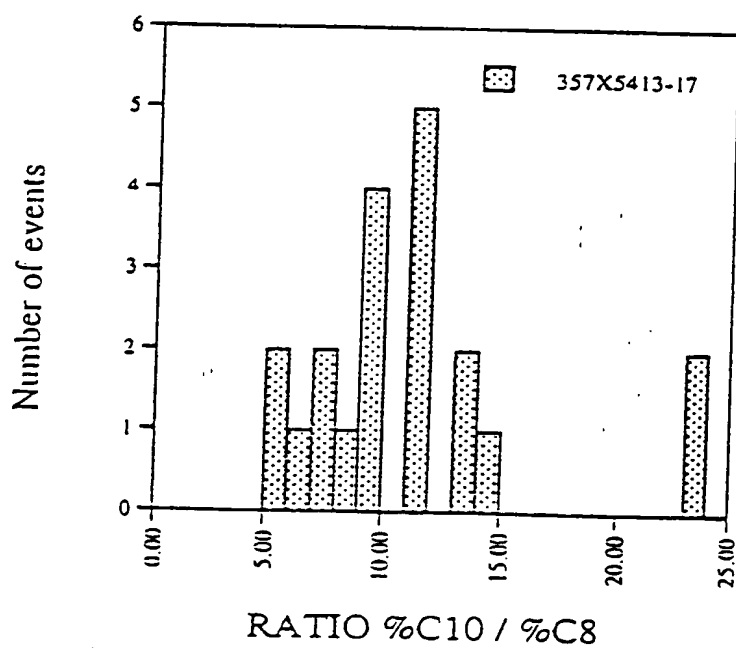


FIGURE 17
1/2

56/66

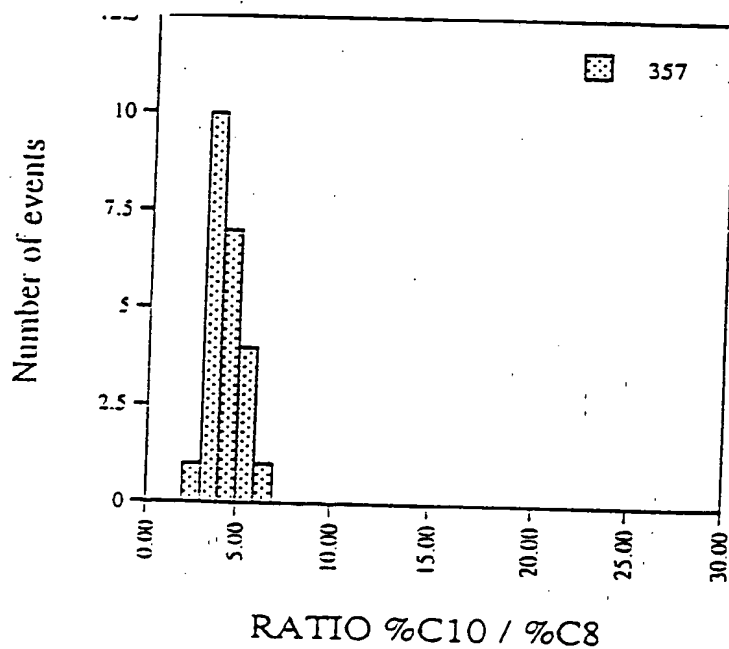


FIGURE 17

2/2

57/66

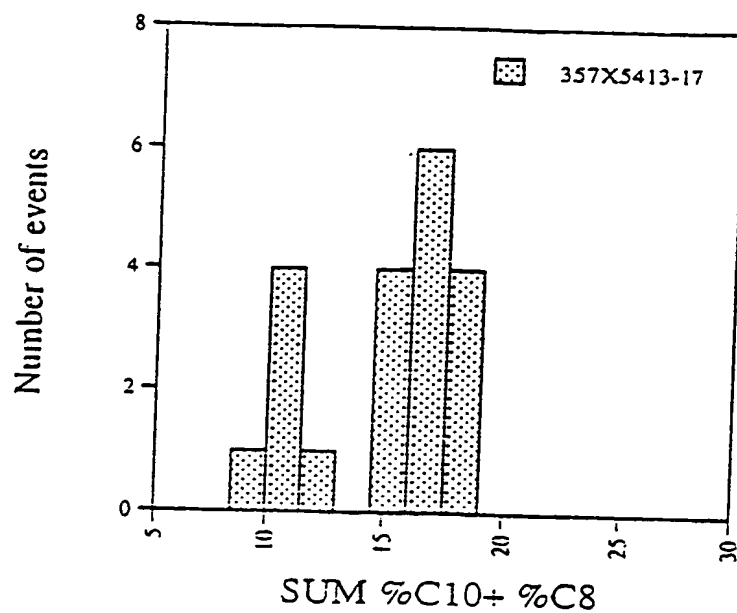


FIGURE 18
1/2

58/66

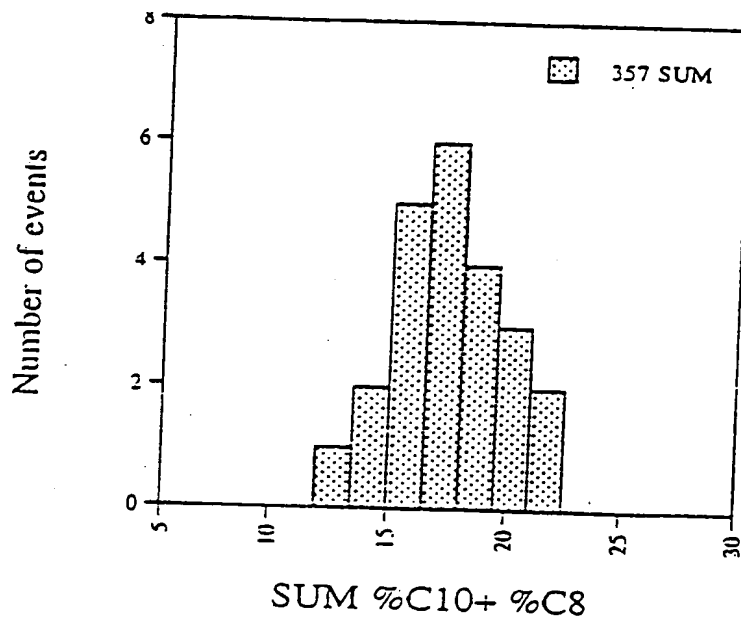
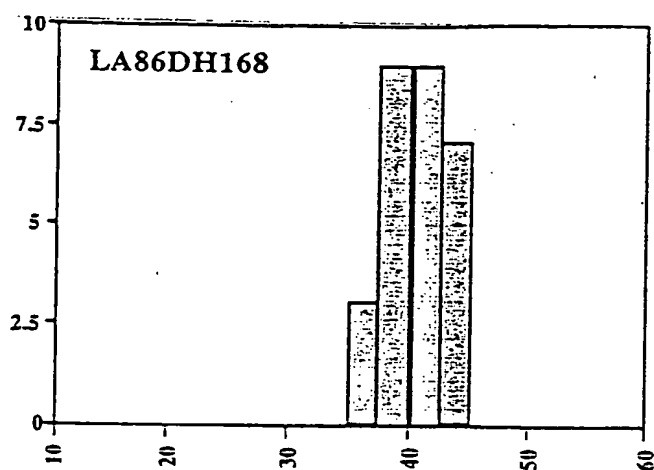


FIGURE 18
2/2

59/66

Number of independent events



12:0 levels (w%)

FIGURE 19
1/3

LD/LL

Number of independent events

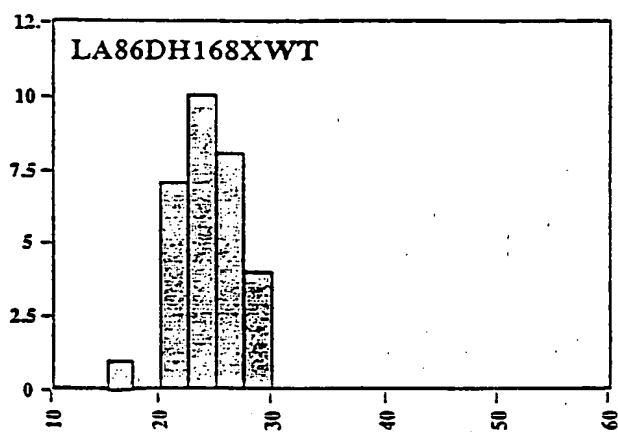


FIGURE 19
3/3

SUBSTITUTE SHEET (RULE 26)

61/66

Number of independent events

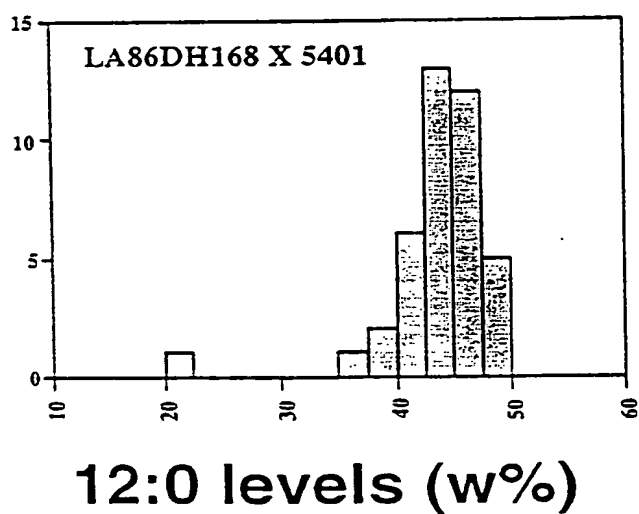
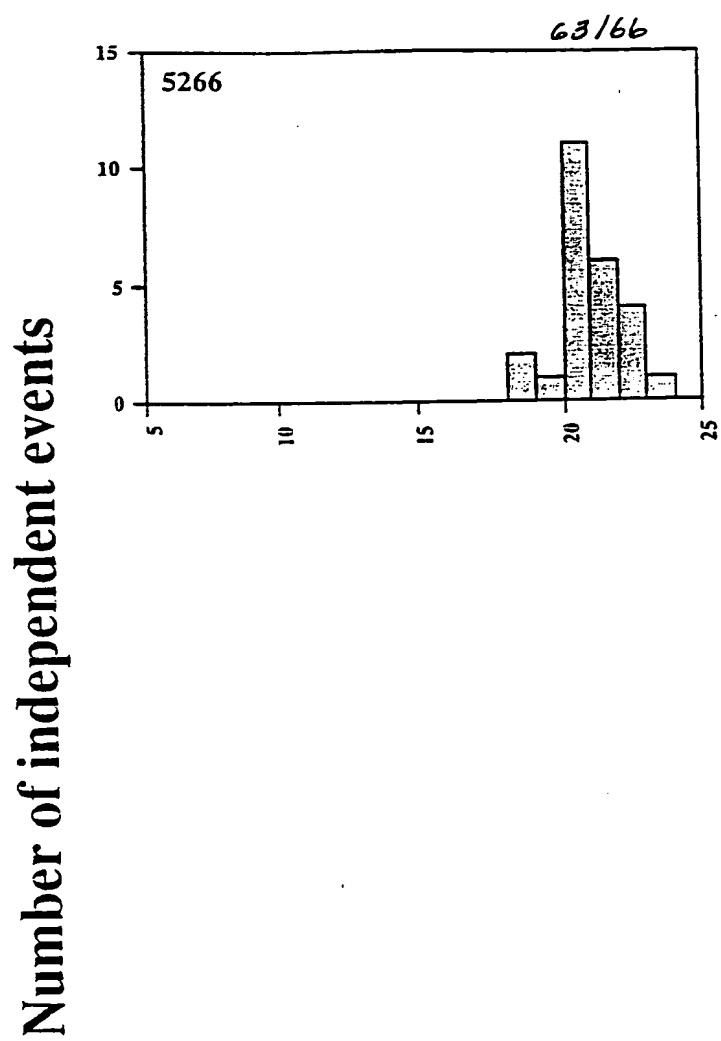


FIGURE 19
2/3

SUBSTITUTE SHEET (RULE 26)

SUBSTITUTE SHEET (RULE 26)

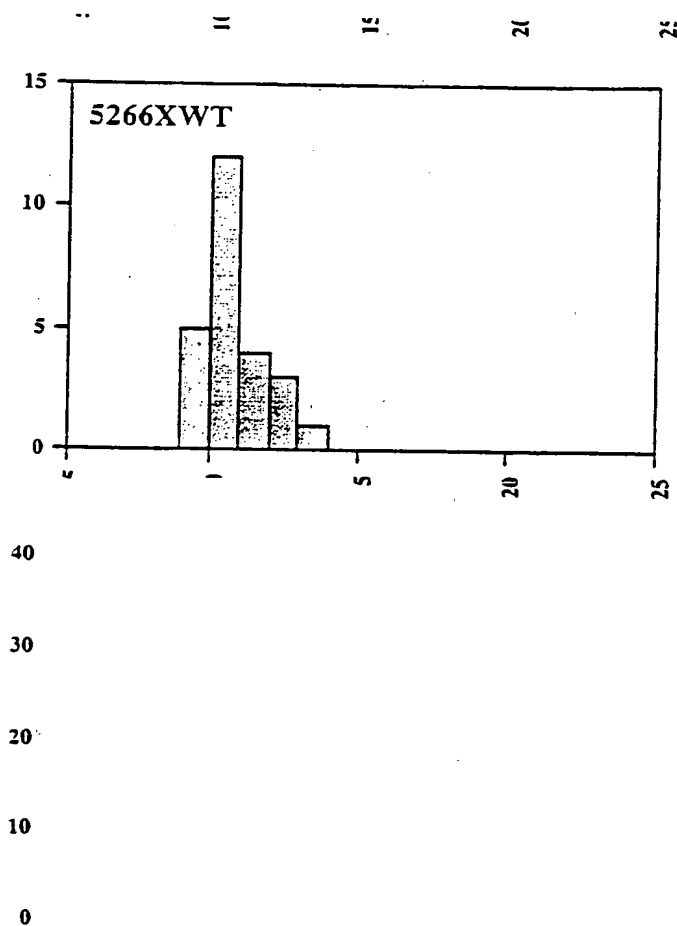


18:0 levels (w%)

FIGURE 21
1/3

64/66

Number of independent events

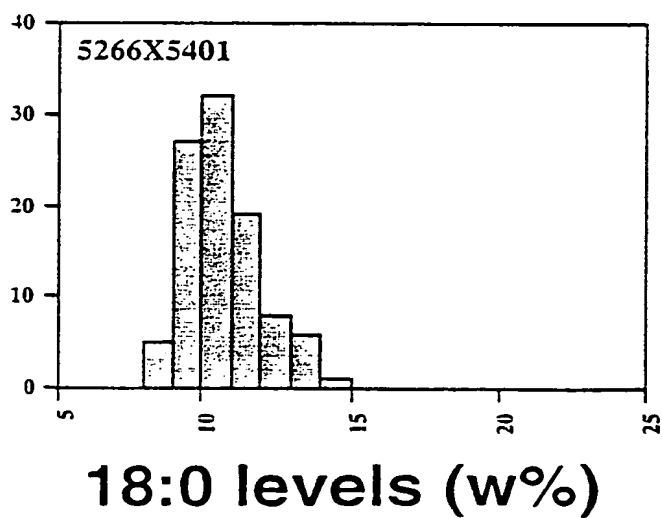


18:0 levels (w%)

FIGURE 21
2/3

C5/66

Number of independent events

FIGURE 21
3/3

66/66

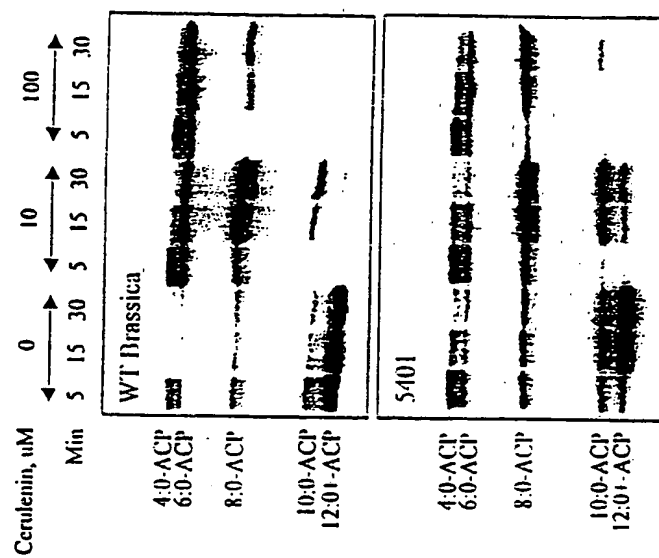


FIGURE 22

THIS PAGE BLANK (USPTO)

This Page is inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images
problems checked, please do not report the
problems to the IFW Image Problem Mailbox**

THIS PAGE BLANK (USPTO)